Институт физиологии им. И. П. Павлова
Российской академии наук

НЕЙРОТЕХНОЛОГИИ

Коллективная монография
Под ред. Ю. Е. Шелепина и В. Н. Чихмана

Санкт-Петербург
ВВМ
2018

УДК 612.84 612.821 617.7
ББК 28.707.3
Н45

Авторы: Алексеенко С. В., Бондарко В. М., Васильев В. Н., Васильев П. П., Жукова О. В., Кезели А. Р., Кирпичников М. П., Ламминпия А. М., Малахова Е. Ю., Малашин Р. О., Моисеенко Г. А., Муравьева С. В., Островский М. А., Пономарев С. В., Пронин С. В., Соколов А. В., Тибилов А. С., Труфанов Г. Е., Фокин В. А., Хараузов А. К., Чихман В. Н., Шелепин Е. Ю. Шелепин К. Ю., Шелепин Ю. Е., Якимова Е. Г.

Н45 **Нейротехнологии** / Под ред. Ю. Е. Шелепина, В. Н. Чихмана / Авторы: С. В. Алексеенко, В. М. Бондарко, В. Н. Васильев, П. П. Васильев, О. В. Жукова, А. Р. Кезели, М. П. Кирпичников, А. М. Ламминпия, Е. Ю. Малахова, Р. О. Малашин, Г. А. Моисеенко, С. В. Муравьева, М. А. Островский, С. В. Пономарев, С. В. Пронин, А. В. Соколов, А. С. Тибилов, Г. Е. Труфанов, В. А. Фокин, А. К. Хараузов, В. Н. Чихман, Е. Ю. Шелепин, К. Ю. Шелепин, Ю. Е. Шелепин, Е. Г. Якимова— СПб.: Изд-во ВВМ, 2018.— 397 с.

ISBN 978-5-9651-1198-5

Нейротехнологии — важнейший раздел цифровых, интеллектуальных технологий, связанных с изучением активности мозга, восприятием и распознаванием сигналов, исследованием и моделированием механизмов принятия решений об изображениях объектов и сцен, анализом «карт» активности нейронных сетей, при планировании и организации целенаправленных действий в реальной и виртуальной среде. Нейротехнологии предопределяют создание систем искусственного интеллекта, конвергенцию человека и автономного интеллектуального искусственного устройства. Новые интеллектуальные технологии, открывают человеку новые возможности и положение в биотехносфере. Нейротехнологии обеспечивают эффективность целенаправленной деятельности, и реабилитацию при психических заболеваниях. В книге рассмотрен широкий круг вопросов нейротехнологий от оптогенетики и искусственных нейронных сетей до управления поведением человека.

Книга предназначена для студентов и аспирантов, психологов, физиологов, врачей и инженеров, создателей систем искусственного интеллекта.

Издание финансируется из гранта РНФ. Номер проекта: 14-15-00918.
Название проекта: «ТЕХНОЛОГИИ ОПТИМИЗАЦИИ И ВОССТАНОВЛЕНИЯ КОГНИТИВНЫХ ФУНКЦИЙ ЧЕЛОВЕКА ВИРТУАЛЬНОЙ СРЕДОЙ».

ISBN 978-5-9651-1198-5

Содержание

Физиологическая оптика и оптогенетика

Глава 1.
Оптогенетика и перспективы восстановления зрения

М. А. Островский, М. П. Кирпичников

Институт биохимической физики им. Н. М. Эмануэля РАН
Институт биоорганический химии им. М. М. Шемякина и Ю. А. Овчинникова РАН
Кафедра молекулярной физиологии и Кафедра биоинженерии Биологического факультета МГУ им. М. В. Ломоносова

Протезирование «слепой» сетчатки

Проблема возможного восстановления приемлемых зрительных функций слепым людям с помощью протезирования сохранившихся в их «слепой» сетчатке нервных клеток чрезвычайно актуальна и близка к решению. Сетчатка «слепая» потому, что в результате той или иной формы дегенеративного заболевания погибают её светочувствительные зрительные клетки — палочки и колбочки. Важно при этом, что нервные клетки, лежащие за зрительными, как правило, остаются здоровыми. Это означает, что если бы тем или иным способом их удалось возбудить, то они могли бы посылать нервные импульсы в мозг, и тогда в определённой степени зрительное восприятия могло бы восстановиться. Такова идея протезирования слепой сетчатки.

Рассматриваются в настоящее время два подхода — электронное и оптогенетическое протезирование. В первом случае множеством электродов возбуждают множество «выходящих» нервных клеток сетчатки, т. е. клеток последнего нервного слоя, посылающих по своим длинным отросткам — аксонам информацию в зрительные области мозга. Во втором случае в сохранившиеся нервные клетки «вставляются» свето-чувствительные белки, и тогда в этих клетках под действием видимого света возникают нервные импульсы. В этом принцип оптогенетического протезирования. Нервные клетки становятся «псевдо-зрительными» клетками. Сейчас приходит понимание того, что оптогенетическое протезирование, несмотря на несомненные успехи электронного протезирования, представляется более перспективным и надёжным [см. обзор: Островский, Кирпичников, 2015].

Дегенеративные заболевания сетчатки глаза как медицинская и острая социальная проблема

Согласно данным Всемирной Организации Здравоохранения примерно 285 миллионов людей страдают нарушениями зрения. Предполагается, что к 2050 году эта цифра удвоится. Дегенеративные заболевания сетчатки, в первую очередь возрастная макулярная дегенерация — основная причина частичной или полной потери зрения у пожилых людей. Проблема это приобретает всё большую остроту, поскольку старение человечества — очевидный демографический феномен. Основные усилия современной офтальмологии направлены на замедление развития дегенеративных заболеваний сетчатки, поскольку действенных способов их профилактики или лечения не существует. Поздние же стадии дегенерации сетчатки — это потеря зрительных клеток, приводящая к полной или частичной слепоте. Оптогенетическое протезирование получило признание как составная часть стратегии генной терапии дегенеративных заболеваний сетчатки [Petit, Khanna, Punzo, 2016].

Оптогенетика как новая методика

Оптогенетика или точнее оптогенетические методы — новое, междисциплинарное, бурно развивающееся направление в биологии, объединяющее физиологию, генную инженерию и оптику. Речь идёт о методике, позволяющей светом регулировать физиологическую активность клетки. Регуляция осуществляется с помощью активируемого светом белка. Ген такого белка доставляется в определённую клетку, например, мозга или сетчатки глаза, и экспрессируется в ней. Белок, а речь идёт о мембранном белке, встраивается в плазматическую мембрану клетки. В зависимости от того, какие именно ионы — катионы или анионы переносит этот белок через мембрану при действии света, мембрана деполяризуется или гиперполяризуется. Деполяризация мембраны приводит к физиологическому возбуждению клетки, её гиперполяризация — к торможению. Светом возбуждается или тормозится активность только той, конкретной клетки мозга или сетчатки, которая содержит светочувствительный белок. Соседняя же клетка, такого белка не содержащая, отвечать на свет не будет. Этим, в частности, оптогенетика принципиально отличает-

ся от воздействия электрического тока или химического вещества, которые возбуждают или тормозят активность сразу большой популяции клеток мозга, сетчатки или другого органа.

С помощью оптогенетики в последние годы получены принципиально новые данные о нейронных сетях, механизмах обучения, памяти, двигательной активности, работе сердца и ряда других функций [см. обзор: Deisseroth, 2011]. Наиболее реальное и близкое медицинское приложение оптогенетических методов в медицине — это офтальмология. Речь идёт об оптогенетическом протезировании дегенеративной сетчатки на поздних стадиях её дегенерации.

Оптогенетика и родопсины

В отличие от генной терапии, задачей которой являются замещение или исправление повреждённого гена, задача оптогенетики — «вставить» в клетку ген светочувствительного белка, который становится «инструментом» управления ею. В настоящее время основными «инструментами» оптогенетики являются светочувствительные мембранные ретиналь-содержащие белки — родопсины [см. обзоры: Островский, 2012; Островский, Фельдман, 2012; Cohen, 2018; Grote et al., 2014; Kaneko et al., 2017; Wiegert et al., 2017]

Возбуждающие и тормозящие канальные родопсины

Для оптогенетики как метода управления физиологической активностью клетки, в первую очередь нервной клетки, принципиально важно иметь в руках два типа «инструментов» — один из которых деполяризует её плазматическую мембрану, т.е. возбуждает клетку, и второй, который её гиперполяризует, в результате чего импульсная активность клетки уменьшается или полностью подавляется.

Деполяризующие «инструменты»

Основной и широко использующийся «инструмент» — это катионный канальный родопсин, а именно канальный родопсин-2 (channelrhodopsin-2 \ChR2\) из зелёной водоросли Chlamydomonas reinhardii. Будучи экспрессирован в нервной клетке, канальный родопсин-2 при действии синего (470 нм) света достаточно быстро (в пределах 50 миллисекунд) деполяризует её. К существенным не-

достаткам канального родопсина-2 относятся его неселективность и относительно малая светочувствительность, что для протезирования сетчатки создает определённые трудности. Тем не менее, он был первым, который был экспрессирован в нейронах дегенеративной сетчатки слепой мыши (трансгенная мышь с пигментным ретинитом) и восстановил её ответы на свет [Bi et al., 2006]

Благодаря успехам современной биоинженерии семейство канальных родопсинов стремительно расширяется и их характеристики улучшаются [см. обзор: Kaneko et al., 2017]. Катион-переносящий канальный родопсин-2 остаётся в большинстве оптогенетических исследований наиболее приемлемым «инструментом» деполяризации клетки, в том числе и в работах, по протезированию дегенеративной сетчатки.

Гиперполяризующие «инструменты»

Гораздо сложнее обстоит дело с гиперполяризующими «инструментами». До последнего времени были известны и использовались в оптогенетических работах следующие «инструменты». Во-первых, самый распространённый из них и широко используемый — это галородопсин. Он представляет собой хлор-переносящий насос, поглощающий в жёлтой области спектра. Поскольку это насос, то он переносит всего один ион хлора на один поглощённый квант света. Это означает, что требуется слишком много света, чтобы достичь приемлемой гиперполяризации. Если при исследовании мозга это не является важным лимитирующим фактором, то для сетчатки использование галородопсина создаёт реальную опасность фотоповреждения. Во-вторых, это свето-активируемые протонные насосы, классическим примером которых может служить бактериородпсин, который был использован и для протезирования сетчатки [Greco et al., 2017]. Однако, в силу целого ряда причин, протонные насосы большого распространения в оптогенетике не получили. Наконец, в-третьих — это семейство не насосов, а именно хлор-переносящих каналов с существенно улучшенными световой чувствительностью и смещёнными в зелёную область спектра характеристиками [Berndt et al., 2016]. Следует подчеркнуть, что эти хлор-переносящие каналы ни что иное, как биоинженерно модифицированные катион-пе-

реносящие каналы родопсина хлорофитовых водорослей. Примерами такого рода модифицированных катион-переносящих каналов могут служить каналы, использованные в работе группы Хагеманна [Wietek et al., 2014] и в работе нашей группы [Долгих и др., 2015]. В этих работах в результате биоинженерной трансформации канального родопсина-2 был получен свето-активируемый анионный канал, т.н. slow ChloC, способный с высокой эффективностью подавлять импульсную активность нейронов в оптогенетических экспериментах.

Недавно был открыт новый класс естественных, анион-переносящих канальных родопсинов (ACR) из криптофитовых одноклеточных водорослей Guillardia theta. [Govorunova et al., 2015, 2017; Li et al., 2016].

Этот канал является наиболее активным из всех известных на сегодняшний день канальных родопсинов, включая катион-переносящий канальный родопсин из хлорофитовых (зелёных) водорослей. Он генерирует исключительно большой (несколько наноампер) фототок, и он намного сильнее тормозит генерацию свето-индуцируемых потенциалов действия — импульсную активность нервной клетки по сравнению с биоинженерно-модифицированным катионным канальным родопсином.

Сравнивая характеристики двух анионных канальных родопсинов — модифицированного катионного канального родопсина и естественного анион-переносящего канального родопсина — мы показали, что свето-активируемый канал анион-переносящего родопсина обладает выраженными преимуществами как по кинетике инактивации, так и по интенсивности торможения импульсной активности нервной клетки, что связано с более негативным значением потенциала реверсии фотоактивируемого им тока, по сравнению с каналом модифицированного катионного канального родопсина [Долгих и др., 2016]. Более того, как мы недавно показали, естественный анион-переносящий канальный родопсин способен, помимо свето-индуцированного торможения импульсной активности нервной клетки, генерировать также при определённых условиях её импульсную активность, преимущественно в аксональных окончаниях. Длительная световая стимуляция приводит к торможению

импульсации, в то время как короткая световая стимуляция, напротив, вызывает генерацию импульсной активности. Это означает, что естественный анион-переносящий канальный родопсин может быть использован в том же самом эксперименте в зависимости от условий световой стимуляции и как эффективный ингибитор физиологической активности клетки и как её активатор [Malyshev et al., 2017]. Биоинженерная модификация естественного анион — переносящего канального родопсина позволяет существенно (более, чем на порядок) увеличить его ион-переносящую активность [Doi, Tsukamoto, Yoshizawa, Sudo, 2017]. Анионные канальные родопсины как новое семейство свето-регулируемых канальных родопсинов позволяют при существенно меньших, чем в случае галородопсина, интенсивностях света гиперполяризовать клеточную мембрану и вызывать торможение импульсной активности клетки. Для протезирования дегенеративной сетчатки это имеет принципиальное значение, поскольку открывает реальные возможности для формирования ON-OFF- и OFF-ON рецептивного поля её ганглиозной клетки.

Мы в настоящее время предпринимаем попытки с помощью катион- и анион-переносящих каналов создать ON-OFF- и OFF-ON рецептивные поля оптогенетически протезированных ганглиозных клеток в сетчатке трансгенной мыши с пигментным ретинитом.

Способы доставки гена в оптогенетике

Исключительно важна в оптогенетике проблема доставки гена светочувствительного белка — канала или насоса в нужную клетку мозга или сетчатки.

В настоящее время известны следующие способы доставки: с помощью вирусов, методом электропорации, с помощью сфокусированного ультразвука и методом лазерной фотопорации. Широко применяется в настоящее время для мозга и сетчатки именно вирусный носитель. Требования к вирусному носителю, особенно имея в виду будущие клинические испытания, следующие: очень низкая мутагенность, очень низкая иммуногенность, нейротропизм, низкий потенциал рекомбинации. Возможны следующие носители — вирусные векторы: адено-ассоциированый вирус, аденовирус, лентивирус, вирус простого герпеса. Адено-ассоциированый вирус (AVV)

является наиболее приемлемым, что обусловлено следующими обстоятельствами:

- он уже используется как вирусный вектор для генной терапии ряда заболеваний;
- он способен инфицировать делящиеся и неделящиеся клетки и эффективно встраиваться в геном хозяина;
- он вызывает слабый иммунный ответ, поскольку не патогенен для человека;
- он не способен реплицироваться, ему для размножения требуется вспомогательный вирус, как правило аденовирус;
- описано несколько серотипов адено-ассоциированого вируса — AAV1-AAV9; они отличаются рецепторами, с которыми связываются. (для сетчатки глаза мыши и обезьяны серотипы разные);
- показана безопасность и эффективность векторов на основе адено-ассоциированного вируса при генной терапии дегенеративного заболевания сетчатки (врожденный амавроз Лебера), что для оптогенетического протезирования сетчатки принципиально важно.

Помимо вируса для успешного оптогенетического протезирования вполне определённых — биполярных или ганглиозных клеток сетчатки, причём разных их типов, принципиально важно иметь эффективный и специфичный промотор. Работы в этом направлении активно ведутся. Как выяснилось, например, для ганглиозных клеток сетчатки глаза трансгенной (пигментный ретинит) мыши эффективность промотора человеческого гамма-синуклеинового гена выше, чем цитомегаловируса [Chaffiol et al., 2017]. По нашим данным, промотор CaMKII, специфичный для глутаматергических нейронов, в сочетании с адено-ассоциированным вирусом эффективно доставляет гены катионного и анионного канальных родопсинов в ганглиозные клетки сетчатки трансгенной мыши.

Таким образом, располагая возможностью доставки с помощью адено-ассоциированного вируса генов катионного и анионного канальных родопсинов в большую популяцию ганглиозных клеток дегенеративной сетчатки, становится реальным создание эффективного ON-OFF и OFF-ON рецептивных полей этих клеток.

Итак, в заключение можно сформулировать основные положения, принципиально важные для оптогенетического протезирования дегенеративной сетчатки глаза:

- Восстановление зрения — первое, наиболее реальное и ближайшее применение методов оптогенетики в клинической практике.
- Ганглиозные клетки дегенеративной сетчатки — наиболее надёжны для оптогенетического протезирования.
- Катионный и анионный канальные родопсины хлорофитовых и криптофитовых одноклеточных водорослей наиболее реальные на сегодняшний день «инструменты» для оптогенетического протезирования ганглиозных клеток.

Насущные задачи:

– оптимизация параметров катионных и анионных канальных родопсинов (светочувствительности, проводимости, времени жизни, спектров поглощения);
– повышение уровня экспрессии генов канальных родопсинов в клетке;
– поиск специфических промоторов для субпопуляций (ON, OFF, ON-OFF) ганглиозных клеток.

- Адено-ассоциированный вирус — единственный приемлемый на сегодняшний день вектор для доставки гена к нервным клеткам сетчатки. Он должен быть сопряжён со специфичным для данного нейрона промотором.
- Вслед за восстановлением монохроматического зрения, используя спектрально отличающееся канальные родопсины, представляется возможным восстановление и цветового зрения.
- Авторы клинических испытаний (компания RetroSense Therapeutics), полагают, что у пациентов, в случае успеха оптогенетического протезирования ганглиозных клеток, может быть восстановлено предметное зрение, и они смогут увидеть предметы (стол и стулья).

Опыты на животных не позволяют оценить качество восстановленного зрения. Однако, успехи в электронном протезировании (система Argus II) свидетельствуют об поистине безграничной пла-

стичности мозга и позволяют надеяться на успех оптогенетического протезирования дегенеративной сетчатки.

Литература

Долгих Д. А., Малышев А. Ю., Саложин С. В., Некрасова О. В., Петровская Л. Е., Рощин М. В., Бородинова А. А., Фельдман Т. Б., Балабан П. М., Кирпичников М. П., Островский М. А. Анионный канальный родопсин, экспрессированный в культуре нейронов и *in vivo* в мозге мыши: светоиндуцированное подавление генерации потенциалов действия. ДАН, 2015. 465. 6. 737–740.

Долгих Д. А., Малышев А. Ю., Рощин М. В., Смирнова Г. Р., Некрасова О. В., Петровская Л. Е., Фельдман Т. Б., Балабан П. М., Кирпичников М. П., Островский М. А. Сравнительная характеристика двух анионных канальных родопсинов и перспективы их применения в оптогенетике // Доклады Академии наук, 2016. Т. 471. № 6. С. 1–4. DOI: 10.7868/S0869565216360238.

Островский М. А. Молекулярная физиология зрительного пигмента родопсина // Биологические мембраны. 2012. № 1–2. С. 38–50.

Островский М. А, Кирпичников М. П. Оптогенетика и зрение // Сенсорные системы, 2015. 25. 4:289–295.

Островский М. А., Фельдман Т. Б. Химия и молекулярная физиология зрения: светочувствительный белок родопсин // Успехи химии. 2012. Т. 81. № 11. С. 1071–1090.

Berndt A. et al. Structural foundations of optogenetics: Determinants of channelrhodopsin ion selectivity. Proceedings of the National Academy of Sciences. 2016 113(4):822–829.

Bi A., Cui J., Ma Y., Olshevskaya E., Pu M., Dizhoor A., Pan Z. Ectopic expression of a microbial-type rhodopsin restores visual responses in mice with photoreceptor degeneration. Neuron, 2006. 50 (1):23–33.

Chaffiol A., Caplette R., Jaillard C., Brazhnikova E., Desrosiers M., Dubus E., Duhamel L., Macé E., Marre O., Benoit P.,

Hantraye P., Bemelmans A. P., Bamberg E., Duebel J., Sahel J. A., Picaud S., Dalkara D. A New Promoter Allows Optogenetic Vision Restoration with Enhanced Sensitivity in Macaque Retina. Mol Ther. 2017. 25(11):2546–2560. doi: 10.1016/j.ymthe.2017.07.011.

Cohen E. D. Retinal Prostheses. In: Webvision: The Organization of the Retina and Visual System. 2018. https://www.ncbi.nlm.nih.gov/books/NBK493746.

Deisseroth K. Optogenetics // Nature Methods, 2011. V. 8, pp. 26–29.

Doi S., Tsukamoto T., Yoshizawa S., Sudo Y. An inhibitory role of Arg-84 in anion channelrhodopsin-2 expressed in Escherichia coli. // Sci. 2017. Rep 7:41879.

Govorunova E. G., Sineshchekov O. A., Janz R., Liu X., Spudich J. L. Natural light-gated anion channels: a family ofmicrobial rhodopsins for advanced optogenetics // Science 349:647–650.

Govorunova E. G., Sineshchekov O. A., Li H., Spudich J. L. Microbial rhodopsins: diversity, mechanisms, and Optogenetic applications. 2017. Annu Rev Biochem 86:845–872.

Greco J. A., et al. Stimulation of Retinal Ganglion Cells Using an Ion-Mediated, Protein-Based Retinal Implant. Investigative Ophthalmology & Visual Science. 2017. 58(8):4184–4184.

Grote M., Engelhard M., Hegemann P. Of ion pumps, sensors and channels — Perspectives on microbial rhodopsins: between science and history // Biochim Biophys Acta. 2014, 1837(5):533–545.

Kaneko A., Inoue K., Kojima K., Kandori Y., Sudo Y. Conversion of microbial rhodopsins: insights into functionally essential elements and rational protein engineering // Biophysical Reviews 2017 9:861–876 https://doi.org/10.1007/s12551–017–0335-x.

Li H., Sineshchekov O. A., Wu G., Spudich J. L. In Vitro Activity of a Purified Natural Anion Channelrhodopsin // J Biol Chem. 2016. 291(49):25319–25325.

Malyshev A.Y., Roshchin M.V., Smirnova G.R., Dolgikh D.A., Balaban P.M., Ostrovsky M.A. Chloride conducting light activated channel GtACR2 can produce both cessation of firing and generation of action potentials in cortical neurons in response to light // Neurosci Lett. 2017 Feb 15;640:76–80. doi: 10.1016/j.neulet.2017.01.026. Epub 2017 Jan 16.

Petit L., Khanna H., Punzo C. Advances in Gene Therapy for Diseases of the Eye // Human gene therapy, 2016.v 27 No 8. pp. 563–579.

Wiegert S., Mahn M., Prigge M., Printz Y., Yizhar O. Silencing Neurons: Tools, Applications, and Experimental Constraints // Neuron 2017 Aug 2; vol. 95(N3): pp. 504–529. doi: 10.1016/ http://dx.doi.org/10.1016/j.neuron.2017.06.050.

Wietek J., Wiegert J.S., Adeishvili N., Schneider F., Watanabe H., Tsunoda S.P., Vogt A., Elstner M., Oertner T.G., Hegemann P. Conversion of channelrhodopsin into a light-gated chloride channel. Science, 2014. 344, 6182, 409–412.

Глава 2.
Сравнение абсолютных порогов чувствительности глаза человека и оптоэлектронной камеры

А. С. Тибилов, В. Н. Васильев

Университет ИТМО, Санкт-Петербург

Сопоставляется абсолютный порог световой чувствительности по яркости наблюдаемых объектов двух «систем», адаптированного к темноте человеческого зрения и современной высокочувствительной камеры, сопряжённой с дисплеем. Если бы задачей было только определение количественного соотношения этих порогов, можно было бы для оценки обратиться к литературным данным или, ещё лучше, провести психофизические эксперименты, предъявляя обеим системам одну и ту же группу стимулов при разных освещённостях (описания такого эксперимента, к сожалению, найти не удалось). Но цель этой статьи иная, аналитическая. И для сумеречно-ночных животных (к которым по соотношению количества палочек и колбочек приближается человек), и для инженеров, занимающихся, например, военной техникой, важно как можно ближе подойти к физическому пределу световой чувствительности. И хотя, как это будет видно из дальнейшего, ни та, ни другая система этого порога не достигают, интересно и поучительно проанализировать детально, как они к этому «пробиваются», сопоставить «стиль» живой природы и техники.

Каждая из систем является трактом из двух частей: в одном случае это сам глаз и зрительные отделы мозга, во втором случае — камера и оператор с пультом и дисплеем. Мы исключим из анализа конечные части обоих трактов («устройств» сложных и недостаточно для количественного моделирования понятных) и будем сравнивать весьма похожие друг на друга структурно и функционально глаз и камеру. Тем более, что основная «битва» с шумами за высокую чувствительность происходит именно здесь, в начальной части тракта. При этом, правда, возникает проблема прямой проверки резуль-

татов расчётов на экспериментальном материале: фундаментальных экспериментальных работ по полю сигналов на выходе ретины в условиях пороговых освещённостей найти не удалось. Следует признать, что в исключённых из анализа конечных частях тракта могут происходить какие-то повышающие чувствительность процессы, в том числе процессы пространственного и временного интегрирования. Поэтому для большей определённости все расчёты привязаны к минимальному времени наблюдения, к времени «кадра», а средства пространственного интегрирования используются только те, которые существуют внутри глаза и камеры.

Предлагаемый материал распадается на две части. В первой части проводится сравнительный анализ приёмников камеры и сетчатки, их шумов.

Во второй части в пределах ограничений, накладываемых шумами, сравниваются уже сами чувствительности глаза и камеры. За счёт простых методов обработки сигнала (интегрирование, дискриминация) возможен «манёвр» по оптимизации совокупности абсолютной чувствительности и набора «тактических» характеристик камеры или глаза: пространственного разрешения, вероятности пропуска пиксела изображения, вероятности «ложного» пиксела. Предполагается, что для глаза эту оптимизацию провёл естественный отбор. Видеокамере мы «предписали» при сравнении тот же набор тактических характеристик. Затем было показано как прохождение сигнала через «конструкцию» сетчатки и камеры обеспечивает заданные вероятности пропуска и ложного сигнала, и рассчитаны пространственное разрешение и пороговые чувствительности для камеры и глаза.

Основная часть выполненных оценок касается преобразования сигнала в ретине и матрице, но важной частью работы являются и выполненные в разделе «Оценка влияния на изображения простых объектов рассеяния света в роговице и в хрусталике глаза» расчёты рассеяния света на хрусталике и роговице, показавшие, что вопреки распространившемуся представлению это рассеяние не снижает чувствительность ночного зрения. Миф о легкомыслии природы, создавшей совершеннейший приёмник ретины при скверной оптике, удалось преодолеть и повысить расчётную пороговую чувствительность сетчатки в несколько раз.

Надо отметить, что в качестве характеристики пороговой чувствительности глаза до сих пор часто используют минимальное количество световой энергии (фотонов), необходимых для того, чтобы заметить точечную вспышку за время одного кадра. Сейчас это дань эпохе возникновения квантовой физики. С практической точки зрения более актуальной характеристикой является используемая в данной статье пороговая наблюдаемая яркость рассматриваемой сцены (ведь глаза и камеры созданы скорее для рассматривания сцен, чем для наблюдения вспыхивающих точек). Эти две характеристики легко пересчитываются друг в друга для камер, а при нынешнем уровне знания устройства ретины и для глаза. Но тривиальность пересчёта не приводит в данном случае к тривиальности его результата: как будет показано в статье, глаз не уступая камерам по пороговой яркости, на порядок уступает им по чувствительности к пороговой световой энергии.

Шумы приёмников

Возьмём за образец одну из лучших серийных камер h-nu 512 фирмы NuVu [http://www.nuvucameras.com/], относящуюся к классу камер на основе охлаждаемых матриц с зарядовой связью с умножением электронов (Electron-multiplying charge-coupled device, EMCCD) — по-прежнему лучшего по чувствительности класса приборов для визуализации мало освещенных сцен.

Эквивалентная шумовая освещенность

Охлаждаемая до –90 ºС матрица 512×512 этой камеры имеет элементарные приемники площадью $s_{\text{М}}$ = (16×16) мкм^2 с квантовым выходом $k_{\text{М}}$ = 0,9, низкий темновой ток (i_{d} = 0,0002 эл/(пиксел с)). Но для принятой нами длительности формирования сигнала $\tau = 0{,}03$ с (обоснование «длительности кадра» приведено ниже) основным источником шумов, ограничивающим чувствительность этого приемника, является специфический для приборов с зарядовой связью CIC-шум (*clock-induced charges*), характеризующийся вероятностью появления шумового электрона при сканировании 1 пиксела ($q = 0{,}001$ эл/пиксел). Переходя к традиционной характеристике класса фотоприемников эквивалентной шумовой освещенности (*equivalent background illumination*, EBI), получаем

$$\mu_M = q/(\tau k_M\, s_M) = 1{,}4\times10^4 \text{ фотонов/(с см}^2).$$

Здесь уместно оценить аналогичную величину для палочек — фоторецепторов адаптированного к темноте глаза. Глаз устроен так, что основным шумом палочек являются так называемые «ложные (или темновые) фотоны» — стандартные электрические импульсные сигналы, принципиально не отличимые от сигнала, формирующегося при поглощении палочкой фотона: процессы их формирования идентичны за исключением первого толчка, связанного у ложного фотона с температурным возбуждением. Поток таких ложных фотонов равен $i_R = 6{,}3\times10^{-3}$ импульса/с на палочку [Baylor, Nunn, Schnapf, 1984]. В той же работе определен квантовый выход палочки $k_R = 0{,}7$. Вопрос о площади, которую надо приписать палочке, чтобы получить соответствующую EBI, решается не так просто, как в случае матрицы. Палочки представляют собой цилиндры диаметром всего 2 мкм и длиной порядка 60 мкм. Они довольно плотно упакованы (с зазором от половины до четверти диаметра) и ориентированы параллельно в направлении распространения света в ретине. Свет даже попав в зазор между двумя палочками, отражаясь от них, будет в конечном счёте поглощён одной из палочек (как в бархате). Поэтому EBI палочек следует считать по формуле $\mu_R = i_R\rho_R/k_R$, где ρ_R — число палочек на единицу площади ретины. В наиболее чувствительной к свету зоне ретины $\rho_R = (1–1{,}4)\times10^7$ палочек/см^2 [Osterberg, 1935], таким образом $\mu_R = 1{,}1\times10^5$ фотонов/(с см^2). То есть EBI ретины, работающей при температуре +36 ºC, всего в 8 раз хуже, чем EBI у матрицы, охлажденной до –90 ºC.

Преимущество фоторецепторов сетчатки

Правда, если сравнивать «темновой ток» палочек не с CIC-шумами, от которых инженеры рано или поздно, наверное, избавятся, а с темновым током приемников матрицы, то соответствующие EBI будут отличаться на три порядка в пользу матрицы EMCCD. Но следует отметить, что это преимущество матрица получает только за счет глубокого охлаждения. Если воспользоваться обычно применяемой для кремниевых приемников оценкой: при снижении температуры приемника на 20 ºC темновой ток падает на порядок, — «кремниевая сетчатка» при температуре тела дала бы EBI порядка

10^8 фотонов/(с см^2). Такая «сетчатка» могла бы иметь полученное выше для палочек значение EBI только при охлаждении приблизительно до –30 ºС. Эта полезная особенность палочек связана с глубоким различием физической природы процессов фотоизомеризации и фотопроводимости, лежащих в основе функционирования рецепторов сетчатки и приемников матрицы соответственно. Различие это иллюстрируется рис. 1. На рис. 1 (справа) представлена двухуровневая молекулярная система, с фотовозбуждения которой квантом энергией $h\nu$ (h – постояная М. Планка, ν — частота) начинается фотоизомеризация и вся последующая цепочка химических и электрохимических превращений, приводящая к формированию нервного импульса. И фотовозбуждение, и тепловое возбуждение происходят здесь с уровня E_0 на уровень E*.

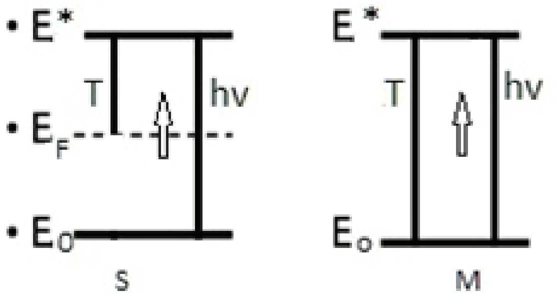


Рис. 1. Преимущество молекулярного приёмника. S — схема уровней полупроводникового приёмника; M — схема уровней молекулярного приёмника; T и hν пути заселения состояния E* тепловым и фотонным возбуждением соответственно; E_F — уровень Ферми

На рис. 1 (слева) представлена схема заселения зоны проводимости в полупроводнике фотовозбуждением. Здесь тепловое возбуждение происходит с уровня Ферми, расположенного у чистого кремния в середине запрещенной зоны. Очевидно, что в случае видимого света тепловые возбуждения будут происходить тут намного чаще (чаще приблизительно в $\exp(h\nu/2k_B T)$ раз, где k_B — постоянная Больцмана, T — температура). К сожалению, полностью это преимущество молекулярного приемника в ретине не используется: в многоатомной молекуле ретинола тепловая фотоизомеризация может

происходить и обходными путями [Ala-Laurila, Donner, Koskelainen, 2004]. Но ретине три порядка по темновому току при прочих равных условиях выиграть у кремниевой матрицы все же удается.

Тепловые шумы

С тепловым движением зарядов связан еще один фундаментальный шум — шум сопротивления при преобразовании токового сигнала в сигнал напряжения. Фактически это шум усиления мощности сигнала к уровню, достаточному для его регистрации. В случае ПЗС он включен в шумы считывания и является препятствием для работы матрицы ПЗС в режиме счета фотонов. Внутреннее усиление в матрицах EMCCD позволяет существенно снизить значения этого шума по сравнению с рассмотренными выше собственными шумами приёмника (во взятом нами за образец приборе среднеквадратическое отклонение этого шума не превышает нескольких процентов от амплитуды фотоэлектрона и эта помеха подавляется практически без ущерба для сигнала). Подобный по природе непрерывный шум генерируется в палочке как тепловой шум мембраны (оболочки палочки), преобразующей флуктуации тока ионов через оболочку в локальные флуктуации разности потенциалов на мембране. Здесь среднеквадратическое отклонение этих флуктуаций достигает уже 25 % от разности потенциалов, генерируемой фотоном [Baylor, Nunn, Schnapf, 1984], и природа предусмотрела механизм его подавления: синапс, соединяющий палочку с расположенной ниже по нервному тракту биполярной клеткой, устроен так, что пропускает импульсы, индуцированные фотонами и «ложными фотонами», и дискриминирует меньшие по амплитуде выбросы теплового шума [Okawa, Sampath, 2007]. Но при этом, как показано ниже, теряется и почти половина фотонов сигнала.

Оценим порог дискриминации шумов в синапсе рецептор — биполяр палочки, необходимый для снижения вероятности ложных сигналов от непрерывной составляющей шумов палочки до значения, на порядок меньшего вероятности генерации ложного фотона. Предполагая нормальное распределение амплитуды шумов и сигналов, и приняв среднюю амплитуду сигнала за единицу, сначала

по вероятности ложного фотона $P_f = 6{,}3\times10^{-3}$ с$^{-1}$ и среднеквадратическому отклонению непрерывных шумов $\sigma_f = 0{,}25$ [Baylor, Nunn, Schnapf, 1984] оценим значение порога дискриминации R из равенства.

$$erfc\left(\frac{R}{\sigma_f}\right) = 0{,}1 p_f \,.$$

В результате оценки $R = 0{,}96$. Зная среднеквадратическое отклонение генерируемого фотоном сигнала $\sigma_s = 0{,}22$ [Baylor, Nunn, Schnapf, 1984] и значение R, можно оценить вероятность K преодоления этого порога дискриминации самим сигналом

$$K = erfc\left(\frac{R-1}{\sigma_S}\right).$$

В результате оценки $K = 0{,}6$, что обеспечивает преодоление барьера в синапсе в среднем 6 из 10 поглощенных сетчаткой при экспозиции фотонов (рис. 2). Фактически это снижает квантовый выход палочки на уровне первого синапса с 0,7 до 0,4 (но при расчете EBI значение квантового выхода следует оставить прежним).

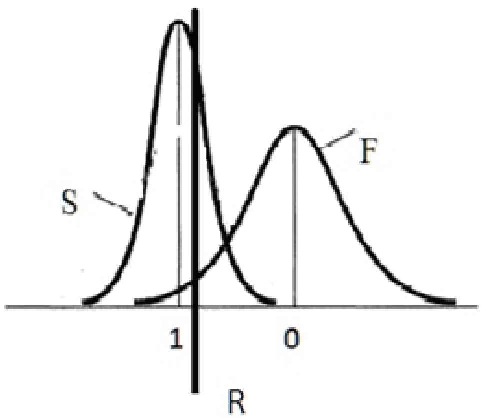


Рис. 2 Подавление тепловых шумов мембраны палочки при амплитудной дискриминации в синапсе палочка — биполяр палочки. F — распределение по амплитуде шумов мембраны; S — распределение по амплитуде фотонных сигналов; R уровень дискриминации

Чувствительность

Эквивалентная шумовая яркость

Освещенность μ имеет тот недостаток, что является смешанной характеристикой, отображающей как свойство объекта, изображенного на матрице (яркость изображения B), так и свойства изображающей оптики (пропускание оптики g, параметр Ω, равный квадрату отношения диаметра действующего отверстия объектива к его фокусному расстоянию); $\mu = Bg\Omega$. В конкретной системе приемник/оптика g и Ω имеют известные значения и в этой системе в качестве характеристики шумов приемника можно ввести понятие эквивалентной шумовой яркости (ЭШЯ), которая рассчитывается как $B_M = (\mu_M/\Omega_M)g_M$ в случае фотоэлектрической матрицы и $B_R = (\mu_R/\Omega_R)g_R$ — в случае ретины, где Ω_M, g_M и Ω_R, g_R — соответствующие характеристики формирующей изображение оптики. Эквивалентная шумовая яркость — это яркость такого объекта, изображение которого, заполняя всю площадь приемника, при используемой оптике создает сигнал, эквивалентный шумовому. Это удобная характеристика, так как сразу дает представление о доступных данному приемнику (и используемой оптике) объектах наблюдения: при яркости таких объектов B_p выполняются условия $B_p > B_M$, $B_p > B_R$.

Значение параметра Ω объектива является обычно компромиссом между качеством изображения и его «энергообеспеченностью». В случае оптики глаза при темновой адаптации зрачка этот компромисс достигнут при значении $\Omega = 0{,}1$ — вполне приемлемом и для широкопольных ночных объективов. В дальнейших расчетах будет использовано это общее значение $\Omega = \Omega_M = \Omega_R$. Значения g_M и g_R будут приняты обычные для хорошей оптики $g_M = g_R = 0{,}9$.

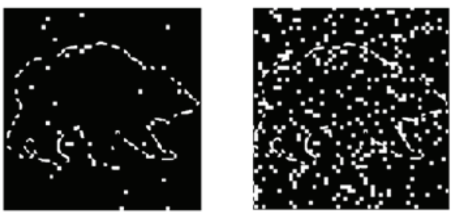
Критерий чувствительности

Ввести понятие ЭШЯ было тем более полезно, что при сравнении матрицы и ретины по чувствительности будут сопоставляться, прежде всего, пороговые яркости доступных им для наблюдения объектов. Правда, само понятие «доступности наблюдения» весьма неоднозначно, находясь зачастую в зависимости даже от смыслового

содержания изображения. Воспользуемся по возможности наиболее простой трактовкой этого понятия.

Чтобы изображения объектов, имеющих во всех частях одинаковую яркость *B*, расположенных на фоне с яркостью, равной нулю, считать доступными наблюдению, должны выполняться два условия:

- Вероятность P_S регистрации фотона на выходе матрицы или на входе зрительного нерва от каждого целиком заполненного изображением элементарного приемника (соответствующего пикселу изображения) за время кадра $\tau = 0{,}03$ с должна быть $P_S > 0{,}68$. Опыт показывает, что простые объекты при отсутствии в их изображении 30 % пикселов распознаются достаточно хорошо. Обоснование выбранной длительности кадра будет приведено ниже.
- Вероятность регистрации фотона в пикселе, соответствующем «изображению» фона, за кадр не должна превышать значения $P_f = 0{,}01$. Психофизические опыты (компьютеризированный вариант Голлин-теста с попиксельно случайным однородным фоном) [Шелепин, Пронин, Шелепин, 2015] показывают, что при приближении к критическому уровню помех, соответствующему 1 % от площади экрана, резко возрастает трудность опознания предлагаемого в тесте разрывно-контурного объекта (см. рис. 3) — фон начинает «запутывать» воображение испытуемого.



a b

Рис. 3 Наблюдение фигуры на фоне случайных шумов.

Вероятность покрытия пикселами контура фигуры — 68 %
а — вероятность ложного фонового пиксела — 0,8 %
б — вероятность ложного фонового пиксела — 12 %

Обеспечение «инсайта» опознания — несомненно, одна из основных задач ночного зрения и следует ожидать, что устройство этого зрения обеспечивает фон не превышающий критический (что является рекомендацией и для устройств ночного электронного зрения). И действительно, такая однопроцентная вероятность фоновых помех проявляется в статистике ложных сигналов в психофизических опытах по определению пороговой энергии наблюдения световых вспышек аккомодированным к темноте глазом [Barlow, 1956; Hecht, Shlaer, Pirenne, 1942].

Используя пуассоновское распределение фотонов и предполагая его таким же у шумов приемников, воспользуемся известной формулой

$$P = \sum_{m=1}^{\infty} \frac{(N)^m}{m!} \exp(-N)$$

вероятности регистрации в каком-либо эксперименте одного или более фотонов при условии, что в среднем в таких экспериментах регистрируется N фотонов, и получим для приемников с площадью S_i следующие выражения:

$$P_{if} = \sum_{m=1}^{\infty} \frac{(k_i \tau \Omega g_i \mathrm{B}_i S_i)^m}{m!} \exp(-k_i \tau \Omega g_i \mathrm{B}_i S_i)\,, \qquad (1)$$

$$P_{iS} = \sum_{m=1}^{\infty} \frac{\left(k_i \tau \Omega g_i S_i \mathrm{B}_i \left(1 + \frac{B_{pi}}{\mathrm{B}_i}\right)\right)^m}{m!} \exp\left(-k_i \tau \Omega g_i S_i \mathrm{B}_i \left(1 + \frac{B_{pi}}{\mathrm{B}_\mathrm{i}}\right)\right), \qquad (2)$$

где i означает «M» в случае фотоэлектрической матрицы и «R» в случае ретины.

Отдельного обсуждения требует используемая в формулах длительность кадра τ. Порядок этой величины известен из ощущения наблюдателем минимального времени между двумя разрешаемыми вспышками низкой интенсивности на темном фоне — 100 мс. Но при исследовании энергетического порога восприятия приходится руководствоваться более строгими критериями: уже в первых психофизических опытах по определению этого порога [Hecht, Shlaer, Pirenne, 1942; Graham, Margaria, 1935] выяснилось, что он, начи-

ная с 10–20 мс, заметно увеличивается с увеличением длительности предъявляемой наблюдателю стимула-вспышки (при 100 мс он уже удваивается). С тех пор в соответствующих опытах используются вспышки, не превышающие по длительности 10 мс, но учитывая необходимость сравнивать глаз с фотоэлектрическими камерами, работающими обычно в телевизионном стандарте, далее используется «компромиссное» значение длительности кадра $\tau = 30$ мс. Подтверждение принятому значению этого параметра можно найти и в современных исследованиях [Herrick, 1973]. При обсуждении результатов расчета этот вопрос будет рассмотрен еще раз.

Численные значения используемых величин сведены в табл. 1.

Таблица 1

Используемые в расчетах параметры EMCCD-камеры и глаза

Параметры	Камера	Глаз
EBI, фотонов/(с см^2)	$1{,}4\times10^4$	$1{,}1\times10^5$
ЭШЯ, фотонов/с см^2 ср	$1{,}6\times10^5$	$1{,}2\times10^6$
Квантовый выход приемника	0,9	0,4
Пропускание оптической системы	0,9	0,9
Площадь элементарного приемника, см^2	$2{,}6\times10^{-6}$ (без биннинга), $4{,}1\times10^{-5}$ (биннинг 4×4)	$1{,}0\times10^{-4}$ (уединенное рецептивное поле), $5{,}0\times10^{-4}$ (переложение 5 рецептивных полей)
Параметр Ω	0,1	0,1
Длительность кадра, с	3×10^{-2}	3×10^{-2}

Не обосновано приведённое в таблице значение коэффициента пропускания оптической системы глаза. Такое обоснование будет дано в специальном параграфе по окончании расчётов (после табл. 2).

Чувствительность фотоэлектронной камеры

Если площадь элементарного приемника задана, уравнение (1) носит проверочный характер, проясняет, можно ли с таким приемником удовлетворить второе из выдвинутых выше условий наблюдаемости объекта. Если же площадь приемника можно проектировать, то приравняв левую часть уравнения к 0,01, его можно использовать

для определения оптимальной S_i. Так подставив, в качестве примера, в правую часть уравнения (1) параметры приемника выбранной нами камеры, получаем значение $P_{Mf} = 9{,}7\times10^{-4}$, с избытком обеспечивающее принятое ограничение на фоновый шум. Приравняв левую часть в уравнении (2) $P_{MS} = 0{,}68$ и решая его относительно B_{pM}, получаем $B_{pM} = 1{,}8\times10^8$ фотонов/(с см^2 ср). Но из уравнения (2) видно, что увеличив площадь приемника до предела, обеспечивающего $P_{Mf} = 0{,}01$ (53×53 мкм), можно получить на порядок меньшее значение предельной яркости $B_{pM} = 1{,}6\times10^7$ фотонов/(с см^2 ср). Во взятой за образец камере к этому идеалу можно приблизиться, воспользовавшись предусмотренным в ней режимом биннинга 4×4 (размер элементарного приемника 64×64 мкм). Размерность матрицы, правда, снизится до 128×128, но размерность «матрицы» глаза при ночном зрении имеет приблизительно такое же значение. Фоновый шум при предлагаемом биннинге чуть выйдет за установленную ему границу ($P_{Mf} = 0{,}014$), пороговая же яркость снизится до $B_{pM} = 1{,}1\times10^7$ фотонов/(с см^2 ср).

Доступный режим биннинга 8×8 выбранной камеры уже неприемлем: в этом режиме $P_{Mf} = 0{,}06$, т.е. далеко выходит за принятый уровень шумов. Для дальнейшего обсуждения устройства ретины здесь будет полезно рассмотреть один из приемов, позволяющих в подобном режиме «сверхбиннинга» все же «неограниченно» увеличивать площадь элементарного приемника и снижать предельную яркость, не выходя за установленные рамки фоновых шумов. Поставим на выходе матрицы дискриминатор, считывающий с ячейки сигнал только в том случае, если он соответствует двум или более электронам в ячейке, тогда уравнения (1) и (2) преобразуются к виду

$$P_{if} = \sum_{m=2}^{\infty} \frac{(k_i \tau \Omega g_i \mathrm{B_i} S_i)^m}{m!} \exp(-k_i \tau \Omega g_i \mathrm{B_i} S_i) \quad (3)$$

,

$$P_{iS} = \sum_{m=2}^{\infty} \frac{\left(k_i \tau \Omega g_i S_i \mathrm{B_i} (1 + \frac{B_{pi}}{\mathrm{B}_i})\right)^m}{m!} \exp\left(-k_i \tau \Omega g_i S_i \mathrm{B}_i (1 + \frac{B_{pi}}{\mathrm{B_i}})\right) \quad (4)$$

.

Из уравнения (3) при отвергнутом ранее режиме биннинга 8×8, подставив соответствующее значение S_M, получаем приемлемый

уровень шумов $P_{Mf} = 1,8\times10^{-3}$. Соответственно из уравнения (4), приравняв $P_{MS} = 0,68$, получаем снижение пороговой яркости в 2 раза — до $B_{pM} = 5,7\times10^{6}$ фотонов/(с см^2 ср). По этому пути можно двигаться и дальше, повышая, например, порог дискриминации до 3, 4, и т.д. электронов и увеличивая площадь элементарного приемника, хотя выигрыш в пороговой яркости при этом становится все меньше и меньше. Кроме того начинают нарастать технические и тактические трудности: увеличение размеров элементарного приемника, матрицы, оптики, охлаждающей системы. Видимо, в этом причина отсутствия в камерах при развитой системе биннинга встроенных дискриминаторов и использования описанного приема как штатного. Но в сетчатку, как можно увидеть дальше, первый, эффективный шаг подобного приема встроен и используется как «штатный».

Чувствительность глаза

Элементарный низкоуровневый фотоприемник ретины

Если элементарным приемником ретины считать минимальный ее элемент, преобразующий попавший на него свет в электрический потенциал, распространяющийся далее от него в виде нервного импульса, то такой приемник — зрительная палочка. Но правильнее, как при биннинге, элементарным приемником считать всю совокупность элементов, формирующих сигнал зрительного нерва, определяющий яркость соответствующего пиксела складывающегося в нашем сознании изображения, т.е. так называемое «рецептивное поле ганглиозной клетки» — фрагмент сетчатки, состоящий из 1–1,5 тысяч палочек. Таким образом, сетчатка напоминает матрицу зрительных палочек в режиме «сверхбиннинга». И действительно, если бы все связанные с ложными фотонами шумовые нервные импульсы палочек рецептивного поля сошлись в соответствующем зрительном нерве, уровень шума был бы слишком велик. Положим среднее число палочек в рецептивном поле одной ганглиозной клетки равным $N = 1,2\times10^{3}$, тогда, воспользовавшись данными, использованными при расчете EBI ретины, получаем площадь такого рецептивного поля $S_R = 10^{-4}$ см^2. Подставив это значение площади и остальные необходимые параметры ретины в формулу (1) получаем $P_{Rf} = 0,12$ (см.

рис. 3b). Это более чем на порядок превышает принятый выше основанный на экспериментальных данных порог приемлемых фоновых шумов. Первое, что в связи с этим можно предположить, — это результат действия обычного встроенного в ретину механизма подавления фонов: разделение каждого рецептивного поля на две части, центральную и периферийную, и последующую компенсацию «прямого» сигнала от первой «непрямым» сигналом второй (за счёт его преобразования в т.н. «горизонтальных связях»). Но этот механизм не может работать при низком уровне фонов: например, при делении рецептивного поля на две равные части вероятность появления ложного фотона в какой-нибудь из них за время кадра становится 0,07, а их появления одновременно в обоих частях — $(0,07)^2$. Т.е. компенсация ложного фотона возможна не более, чем в 5 случаях из 100. Такой механизм остаётся бессмысленным и при освещённостях, соответствующих пороговым интенсивностям (легко показать, что вероятность компенсации равна в этом случае 0,1). Более того, он становился бы вредным, сокращая в 2 раза полезную площадь рецептивного поля, но исследования показали [Peichl, Wässle, 1983; Jensen, 1991], что при темновой адаптации относительная площадь периферии и сигналы от неё резко сокращаются и «прямой» сигнал формирует всё рецептивное поле.

Можно предположить, что снижение шумов — результат (как и в рассмотренном выше освоении «сверхбиннинга» в фотоэлектрической матрице) действия амплитудного дискриминатора. Но положение такого дискриминатора в ретине долго не удавалось найти.

Формирование сигнала, приходящего в основание зрительного нерва, проиллюстрировано рис. 4.

От палочек до зрительного нерва, сходясь, сигналы проходят через 4 слоя синапсов [Sterling, Freed, Smith, 1988]. Свою функцию предполагаемый дискриминатор может выполнить, располагаясь в II, III или IV слое синапсов. Анализ данных психофизических исследований [Тибилов, Шелепин, 2013] указывал на III слой, как наиболее вероятное место его расположения. Действительно, синапс III слоя (между амакриновой клеткой A11 и аксоном биполяра колбочек), как выяснилась недавно [Ala-Laurila, Rieke, 2014], в условиях темновой адаптации глаза и предельно низких интенсивностей ока-

зался дискриминатором нервных импульсов по амплитуде с порогом, по-видимому, «×2». То есть импульс от одного фотона, поступающий на амакриновую клетку от биполяра палочек, не проходит через этот синапс, но суммарный сигнал от 2 и более фотонов — проходит. Это открытие имеет важное значение для понимания механизма реализации высокой чувствительности глаза.

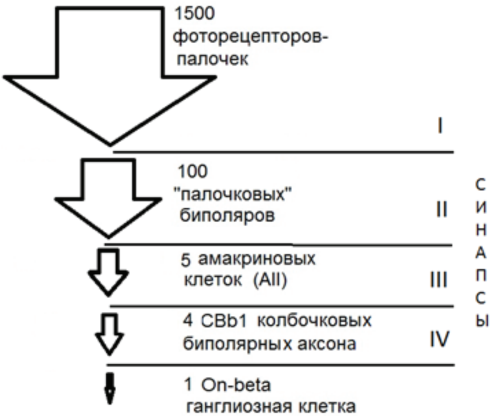


Рис. 4. Схема конвергенции сигналов в палочковой составляющей ретины

Воспользуемся справедливой для пуассоновского распределения общей формулой

$$P_n = 1 - \left[1 + \sum_{m=1}^{n-1} \frac{\left(N \frac{s}{S} \right)^m}{m!} \right]^{\frac{S}{s}} \exp(-N) \qquad (5)$$

вероятности регистрации n и более фотонов в одной из одинаковых ячеек с площадью s в эксперименте, в котором на суммарную площадь всех ячеек S в среднем падает N фотонов.

Так как возбуждение в зрительном нерве возникает, стоит двум или более шумовым или сигнальным нервным импульсам пройти хотя бы через один из пяти синапсов амакриновая клетка A11 — аксоны биполяра колбочек — вероятности таких событий будут иметь вид

$$P_{Rf} = 1 - \left[1 + k_R \tau \Omega g_R \mathrm{B}_R s_R\right]^{\frac{S_R}{s_R}} \exp(-k_R \tau \Omega g_R \mathrm{B}_R S_R), \quad (6)$$

$$P_{Rs} = 1 - \left[1 + k_R \tau \Omega g_R \mathrm{B}_R \left(1 + \frac{B_{pR}}{B_R}\right) s_R\right]^{\frac{S_R}{s_R}} \exp\left[-k_R \tau \Omega g_R \mathrm{B}_R \left(1 + \frac{B_{pR}}{B_R}\right) S_R\right]. \quad (7)$$

По формуле (6) получаем приемлемое значение $P_{\mathrm{Rf}} = 1{,}7\times10^{-3}$, по (7) — пороговую яркость $B_{\mathrm{pR}} = 3{,}7\times10^{7}$ фотонов/(с см^2 ср), приближающуюся к пороговым яркостям выбранной за образец камеры.

Описанные здесь элементарные приемники на основе «уединенного» рецептивного поля ганглиозной клетки существуют, но только на периферии основного массива палочек.

Элементарный приемник на основе переложения рецептивных полей

С точки зрения ограничения, наложенного нами на фоновые шумы, чувствительность элементарного приемника может быть еще увеличена либо за счет увеличения рецептивного поля отдельной ганглиозной клетки, либо за счет сложения сигналов нескольких соседних ганглиозных клеток. Действительно, полученное значение $P_{\mathrm{Rf}} = 1{,}7\times10^{-3}$ в несколько раз меньше ограничивающего предела. Объединим, например, сигналы от 5 соседних рецептивных полей (т.е. 25 амакриновых клеток): вероятность фонового шума согласно выражению (6) останется приемлемой — $P_{\mathrm{Rf}} = 8\,10^{-3}$ (см. рис. 3а), а пороговая яркость согласно выражению (7) снизится в 2,6 раза и станет $B_{\mathrm{pR}} = 1{,}4\times10^{7}$ фотонов/(с см^2 ср). То есть при таком объединении пороговая яркость адаптированного к темноте глаза была бы всего в 1,3 раза больше пороговой яркости выбранной EMCCD-камеры в режиме 4×4 биннинга.

Может показаться удивительным, но этот «ход» действительно реализован в сетчатке и при этом достаточно оригинальным способом: рецептивные поля не растут, сигналы от них не складываются, а эффект достигается пространственным перемешиванием в режиме темновой адаптации нескольких соседних рецептивных полей [DeVries, Baylor, 1997], исключающим пространственное разрешение в пределах участка сетчатки, занятого

рецептивными полями входящих в этот пул ганглиозных клеток. Разница с традиционным способом интегрирования — сложением сигналов соседних пикселов, проявится в формируемом изображении: вспышки от пороговых пар фотонов не будут синхронно возникать в этих пикселах, а будут блуждать между ними случайным образом. Такими перемешанными рецептивными полями покрыта бóльшая часть выделенной под палочки «ночной» поверхности ретины.

Наблюдаемая в устройстве «ночной» ретины тенденция повышения чувствительности за счет увеличения площади элементарного приемника приводит в конечном итоге к резкому ухудшению углового разрешения ночного зрения по сравнению с дневным. В перемешанном рецептивном поле пяти ганглиозных клеток размещается порядка 6000 палочек и угловой размер пиксела становится в рамках редуцированной оптической схемы глаза $\alpha = 46'$, что соответствует наблюдаемому на практике угловому разрешению 40–50′ при ночном зрении [Глезер, Подвигин, 1971; Barlow, I958]. Рассчитанные значения минимальных наблюдаемых яркостей во всех рассмотренных случаях сведены в табл. 2.

Таблица 2

Сравнение расчетной минимальной яркости, наблюдаемой глазом и EMCCD-камерой

Параметры	Камера без биннинга	Камера с биннингом 4×4	Сетчатка из не перекрытых уединённых рецептивных полей	Переложение 5 рецептивных полей
Яркость минимальная, фотоны/(с см² ср), (мкнт)	$1{,}8\times10^{8}$ (196)	$1{,}1\times10^{7}$ (12)	$3{,}7\times10^{7}$ (40)	$1{,}4\times10^{7}$ (15)
Энергетический порог (на зрачке) возбуждения пиксела изображения за кадр, фотоны	1,4	1,3	11	21
Вероятность ложного сигнала в пикселе за кадр с $\tau = 30$ мс	$9{,}7\times10^{-4}$	$1{,}4\times10^{-2}$	$1{,}7\times10^{-3}$	8×10^{-3}

Оценка влияния на изображения простых объектов рассеяния света в роговице и в хрусталике глаза

Оценка энергетических потерь на рассеяние при изображении на сетчатке точечного источника света

Везде при дальнейших оценках будет использоваться модель глаза с сетчаткой, представляющей собой шаровой сегмент с центром кривизны в формирующем изображение и одновременно рассеивающем зрачке. К этой конструкции привязывается соответствующая сферическая система координат с осью, совпадающей с оптической осью глаза. Для упрощения выкладок радиус кривизны сетчатки полагается равным 1. Тогда в силу симметрии доля $\delta I/I$ попадающего в глаз потока света I от помещенного на оптической оси точечного источника, рассеиваемая в зрачке под углом θ к оси в пределах от $\theta + d\theta/2$ до $\theta - d\theta/2$, равна

$$\frac{\delta I}{I} = \eta(\theta)\sin\theta d\theta\,,$$

где $\eta(\theta)$ — хорошо исследованная для глаза индикатриса рассеяния [Vos Johannes, 2003]. Для оценок ее можно экстраполировать двумя функциями

$\eta_1 = 0{,}138\theta^{-1{,}8}$ при $\theta < 1{,}75\times10^{-3}$ рад (т.е. меньше 6′),

$\eta_2 = 1{,}5\times10^{-3}\theta^{-2} + 5\times10^{-5}\theta^{-3}$ при $\theta > 1{,}75\times10^{-3}$ рад (т.е. больше 6′).

Для оценки полной энергии рассеянного света I_d можно воспользоваться выражением

$$I_d = 2\pi I\left[\left(\int_{\theta_{min}}^{0{,}00175}\eta_1(\theta)\sin\theta d\theta\right)+\left(\int_{0{,}00175}^{\pi}\eta_1(\theta)\sin\theta d\theta\right)\right]. \qquad (1п)$$

В качестве минимального угла θ_{min} естественно взять половину дифракционного угла, которая при умеренных освещенностях (диаметр зрачка 3 мм) и длине волны $\lambda = 0{,}5$ мкм равна 20». Рассеяние при расчетах резко падает с ростом θ, поэтому во втором интеграле верхним пределом можно положить $\pi/6$ (при уменьшении верхнего предела с 30° до 25° значение I_d уменьшается всего на тысячную). При этих условиях оказывается $I_d = 0{,}77I$, причем

свет, рассеянный в углы меньше 6′ составляет 0,54I, а в бóльшие углы, соответственно, 0,23I. Полученное значение I_d соответствует общепринятому представлению о рассеянии в зрачке 75 % попадающего в глаз света.

Но следует обратить внимание на то, что рассеяние в основном перераспределяет световые потоки в угловых пределах, малых по сравнению с интересующим нас угловым разрешением ночного зрения. Поставив верхним пределом во втором интеграле формулы (1п) половину углового размера элементарного приемника аккомодированного к темноте глаза (23′), получаем, что за этот предел рассеивается только 9 % света, поступающего в глаз от точечного источника.

Оценка параметров ореола вокруг изображения круга на сетчатке, возникающего из-за рассеяния света на роговице и хрусталике глаза

Представим изображение на сетчатке круга с центром на оптической оси глаза, угловым размером 2J и яркостью B в виде совокупности малых площадок, положение которых определяется в принятой системе координат зенитным и азимутальным углами i и α. Интенсивность света, формирующего изображение этого элемента, равна

$$\delta I_{i\alpha} = \frac{I \sin i\, di\, d\alpha}{\int_0^J di \sin i \int_0^{2\pi} d\alpha},$$

где I — интенсивность света, формирующего изображение круга.

Параллельный пучок света, распространяющийся вдоль оси и формирующий центральную точку круга, проходя через зрачок, формирует на сетчатке симметричное поле рассеянного света. Представим его в виде совокупности вложенных друг в друга усеченных шаровых сегментов. Выберем один из них, характеризуемый зенитным углом его положения θ и угловой шириной $d\theta$. Каждый из формирующих изображение круга пучков образует такое же сегментированное поле рассеянного света, как и центральный пучок, но наклоненное и смещенное относительно центрального поля. Часть

сегментов этого наклонного поля (соответствующих углам рассеяния от ($\theta - i$) до ($\theta + i$)) пересечет выбранный центральный сегмент θ, и в местах пересечения интенсивность рассеянного из наклонного пучка света создаст вклад в интенсивность рассеянного света в сегменте θ. Рассмотрев геометрию этих пересечений можно показать, что отношение этого вклада к полной интенсивности света, распространяющегося в наклонном сегменте, зависит только от i. Тогда весь вклад δI_θ рассеянного из формирующих изображение круга пучков света в освещение сегмента θ можно получить, умножив суммарную интенсивность рассеянного света от каждого из этих пучков (для $(\theta - J) > 6'$ — интересующей нас области внешнего ореола вблизи границы изображения),

$$\delta I_{i\alpha} \int_{\theta-i}^{\theta+i} \left(1{,}5\frac{10^{-3}}{x^2} + 5\frac{10^{-5}}{x^3} \right) 2\pi \sin x dx$$

на отношение площади сегмента θ к суммарной площади наклонных сегментов этого пучка

$$\frac{2\pi \sin\theta d\theta}{4\pi\theta i},$$

и проинтегрировав полученное произведение по всем пучкам

$$\delta I_\theta = I \frac{d\theta \sin\theta}{\theta \int_0^J di \sin i} \int_0^J di \frac{\sin i}{2i} \int_{\theta-i}^{\theta+i} dx \left(1{,}5\frac{10^{-3}}{x^2} + 5\frac{10^{-5}}{x^3} \right) 2\pi \sin x . \qquad (2п)$$

Учитывая, что

$$\delta I_\theta = B_\theta s 2\pi \sin\theta d\theta , \ I = Bs2\pi \int_0^J di \sin i ,$$

где s — площадь зрачка, получаем отношение яркости ореола к яркости изображения круга на сетчатке

$$\frac{B_\theta}{B} = \frac{\pi}{\theta} \int_0^J di \frac{\sin i}{i} \int_{\theta-i}^{\theta+i} dx \left(1{,}5\frac{10^{-3}}{x^2} + 5\frac{10^{-5}}{x^3} \right) \sin x .$$

Взяв круг интересующего нас углового размера $2J = 46'$ (размер элементарного приемника при ночном зрении, см. основную ста-

тью), получаем, что яркость ореола на угловом расстоянии 6′ от края изображения составляет всего 2,3 % от яркости самого изображения (с учетом снижения яркости изображения из-за рассеяния). На угловом расстоянии в $2J$ яркость ореола снижается уже до 0,8 % яркости изображения (т.е. становится меньше предполагаемой эквивалентной яркости собственных шумов ретины, см. выше).

Проинтегрировав в формуле (2п) правую часть по всем θ от $(J+6')$ до $\pi/6$, получим оценку уменьшения яркости изображения круга за счет потерь на рассеяние — 8,1 %. Таким образом, принятое для зрения с предельным разрешением значение потерь на рассеяние в 75 % на ночное зрение не распространяется: при ночном зрении «пропускание» оптической системы глаза можно положить равным 0,9–0,8, т.е. порядка потерь в объективах приборов.

Сделанная оценка позволяет считать влияние рассеяния на формирование на сетчатке «пороговых» изображений слабым, а расчеты, выполненные выше без учета рассеяния, достоверными.

Сравнение с экспериментом

Теперь следует рассмотреть вопрос о том насколько приведённые в табл. 2 величины соответствуют существующим экспериментальным данным.

Основные экспериментальные работы по определению пороговой яркости для глаза были выполнены в середине прошлого века и развивались в двух направлениях. В первой группе работ ставилась цель выявить минимальную наблюдаемую яркость объектов. Для этого приходилось увеличивать время наблюдения и угловые размеры объектов. В качестве примера достаточно привести результат одного из таких исследований [Denton, Pirenne, I954], часто упоминаемого в связи с обнаруженной в нем уникально низкой наблюдаемой пороговой яркостью — $1{,}5\times10^5$ фотонов/(с см^2 ср), или 0,2 мкнт. Результат этот получен при наблюдении стимула с угловым размером 42° при времени «рассматривания» 5 с. Очевидно, что такое значительное (почти на 2 порядка) уменьшение пороговой яркости относительно той же величины, рассчитанной здесь для элементарного приёмника при наблюдении 1 кадра, связано с пространственно-временным усреднением по потоку кадров. Со-

гласно принятой здесь модели такое интегрирование должно происходить вне ретины, т.е. в зрительных отделах мозга. Разумеется, такая возможность — это ещё одно замечательное свойство зрения, но к предпринимаемому здесь сравнению глаза и камеры оно имеет косвенное отношение: подобное интегрирование можно проделать, повысив чувствительность за счёт пространственно-временного разрешения, и с кадрами на выходе камеры. Расчёт показывает, что снизив в выбранной нами камере угловое разрешение до 10 угловых градусов и накапливая сигнал в течение 5 сек, предъявляемое в опытах Дентона и Перена изображение круга мы получили бы на экране дисплея как изображение с контрастом 14 % относительно шумового фона при флуктуациях этого фона порядка 1 %. Изображение на хорошем дисплее наблюдаемое. Такого рода обработка не может нарушить обнаруженного здесь паритета чувствительности камеры и глаза: как почти всегда бывает с датчиками, чувствительность их определяется соотношением сигнал — шум на первом каскаде преобразователя, в нашем случае в кадре на выходе матрицы или ретины.

Возвращаясь к экспериментам, во второй группе работ ставилась цель выявить минимальную порцию световой энергии, вызывающую зрительное ощущение. Для этого приходилось, наоборот, уменьшать угловые размеры объектов и время наблюдения (до величины не превышающей длительности кадра!). Хотя пространственно-временные масштабы в большинстве выполненных по этому направлению экспериментов меньше используемых в представленных здесь расчетах, но сопоставление результатов возможно благодаря действующим на этих масштабах законам Рикко и Блоха–Шарпантье. (Это эмпирические законы, отражающие тот факт, что при уменьшении размера или длительности стимула ниже некоторого предела пороговая энергия обнаружения светового импульса не меняется). Прежде всего, из этого сопоставления следует, что принятые в данной работе значения длительности кадра $\tau = 0{,}03$ с и углового размера элементарного приемника $\alpha = 46'$ находятся в согласии с результатами экспериментов. (Соответствующие величины трактуются в экспериментах как наибольшие значения углового размера и длительности стимула, при которых еще выполняются законы Рикко и Блоха–Шарпантье соответ-

ственно [Глезер, Подвигин,1971; Barlow, I958].)

Сложнее разобраться с вопросом о соответствии экспериментальным данным рассчитанной минимальной наблюдаемой яркости. Минимальная яркость в нашем случае связана минимальной порцией фотонов *n* на входе зрачка, приходящейся на элементарный приемник и вызывающей зрительное ощущение в месте его расположения, соотношением

$$B_{pR} = \frac{n}{\tau S_R \Omega} \tag{8}$$

Но величина *n*, чаще всего определяемая в экспериментах этого направления, до сих пор не имеет общепризнанного значения: даже в наборе «классических» экспериментов, упоминаемых обычно в обзорах. Эта величина, полученная разными группами, меняется в несколько раз от значений n = 25–30 [Bouman, Van der Velden, 1947; Bouman, Van der Velden, 1948], близких к здесь рассчитанным (обратный счет по формуле (8) от вычисленной здесь минимальной яркости дает n = 21), до n = 90–100 [Barlow, 1956; Hecht, Shlaer, Pirenne, 1942]. Велик разброс и в сериях экспериментов одной группы.

В рамках предлагаемой модели причина такой неопределенности ясна: если при яркости, много меньшей пороговой, редко наблюдаемые сигналы от элементарного приемника были связаны с возбуждением 1 или 2 ганглиозных клеток, входящих в пул приемника, с ростом яркости стимула, предлагаемого наблюдателю, тот все чаще наблюдает яркие вспышки одновременно от 3, 4, и даже 5 ганглиозных клеток. Опыт психофизических исследований показывает, что наблюдатель склонен считать яркие вспышки, пропуская слабые [Barlow, 1956], отчего порог смещается в сторону бóльших яркостей (и, соответствен, бóльших n). Действительно, специальные инструкции [Sakitt, 1972] наблюдателю помогали снизить порог в 2–3 раза, но к времени этого исследования основной экспериментальный материал в этой области был уже собран. В рамках данной модели был проведен расчет порогового B_{pR} при условии, что наблюдатель будет засчитывать только самые яркие вспышки с участием всех 5 ганглиозных клеток пула элементарного приемника. При этом получилась

пороговая яркость соответствующая $n = 95$.

При таком состоянии экспериментального материала приходиться довольствоваться тем, что, как и следовало ожидать, рассчитанный порог наблюдаемости оказался нижней границей совокупности его измеренных значений.

Наблюдение объектов с угловым размером менее 46′. «Ночезрительные» трубы

Результат таких наблюдений зависит от степени «перемешанности» рецепторов внутри области элементарного приемника. Многократное подтверждение закона Рикко [Barlow, I958] показывает, что в масштабе от 6′ до 40′ они совершенно перемешаны, т. е. в этих пределах для возбуждения пиксела за время кадра на элементарный приемник должно придти одно и то же число фотонов, независимо от того, какую часть элементарного приемника займет такое малоугловое изображение. Тогда очевидно, что пороговые яркости будут расти пропорционально отношению площади элементарного приемника к площади изображения, меньшего, чем элементарный приемник. С этим явлением было связано применение на флоте до появления радаров так называемых «ночезрительных» подзорных труб.

Раннее обнаружение парусным военным кораблем противника было важным тактическим преимуществом. На море довольно часто возникают такие благоприятные метеоусловия и условия освещения, что в дневное время с мостика можно обнаружить парусный корабль по появлению из-за горизонта только верхней части его парусов. Это соответствует расстоянию приблизительно в 30 км, и парус представляется пятном с угловым размером порядка 4–6′. В звездную, но безлунную ночь парус имеет яркость приблизительно 50 мкнт, превышающую рассчитанную здесь яркость, предельно низкую для глаза. Но из-за углового размера, в 10 раз меньшего углового размера элементарного ночного приемника ретины, пороговая яркость этого паруса станет в 100 раз больше истинной яркости паруса и глазом он останется невидимым. Если же для наблюдения за горизонтом применить подзорную трубу с десятикратным или бóльшим увеличением, обеспечивающим изображение паруса на сетчатке с угловым

размером 46′ и более, яркость изображения перейдет границу пороговой яркости и парус станет видимым. Очевидно, что в описанной процедуре объектив трубы должен быть не менее 8–10 см в диаметре. Польза от «ночезрительной» трубы распространяется и на другие ситуации: например, обнаружение пенных бурунов на рифах. По преданиям на Титанике впередсмотрящий то ли не был адаптирован к темноте, так как вышел из яркого помещения то ли был заперт бинокль и ключ утерян.

Заключение

Выполненный расчет показал, что в первичных (сформированных на выходе матрицы или ретины за 1 кадр с максимальным разрешением) изображениях чувствительность к предельно низкой яркости объектов невооруженного, но полностью адаптированного к темноте глаза, не ниже чувствительности глаза, вооруженного современной оптико-электронной камерой с охлаждаемой кремниевой матрицей. Отсутствие подавляющего шумы охлаждения компенсируется в глазе использованием пока не существующего в технике типа приемника — молекулярного.

Важную роль в этом паритете сыграло также превышение на порядок площади элементарного ночного приемника глаза над площадью элементарного приемника типовой кремниевой матрицы при одинаковой, свойственной ночному зрению размерности поля зрения (близкой к стандарту 128×128 пикселов) и одинаковом уровне шумов. Не в последнюю очередь это преимущество реализуется в ретине без увеличения шума за счет использования в каждом элементарном приемнике ретины нелинейных элементов, подавляющих шумы.

Показано, что с природы можно снять обвинение в нелепом устройстве глаза: сочетании очень чувствительного приемника — ретины и плохой, сильно рассеивающей оптики. Справедливое, вообще-то, в случае дневного зрения высокого разрешения представление о больших энергетических потерях из-за рассеяния света в роговице и хрусталике на ночное зрение не распространяется: из-за узкой индикатрисы рассеяния 90 % рассеянного света остается в пределах 46′ — углового размера элементарного приемника «ноч-

ной» ретины.

Выявленный паритет аккомодированного к темноте зрения и оптико-электронных камер довольно устойчив. Радикально, например, на порядок, повысить пороговую яркость камеры можно, только увеличив на порядок площадь элементарного приемника, и, соответственно, чтобы не уменьшить поле зрения, увеличив размер матрицы. Правда, для этого, при сохранении высоких требований к фоновым шумам, необходимо на порядок понизить EBI. Все эти возможности есть уже сейчас (составные матрицы, усиленный биннинг, охлаждение фотокатода в матрице EBCCD (*Electron bombardment charge coupled devices*) и пр.), но при этом придется увеличивать на порядок мощность криогенной системы и габариты оптической системы (чтобы сохранить светосилу и угловое разрешение). Частично (увеличение площади элементарных приемников и матрицы) описанная программа уже реализована в больших специальных камерах, например, для обсерваторий, но вряд ли их стоит сравнивать с глазом.

На практике из-за неудобства процедуры темновой адаптации глаза (до часа предварительной выдержки в темноте, неустойчивость к засветкам) при деятельности в условиях низкой освещенности предпочтение отдается камерам и подобным им устройствам. Выстояв в сложной конкуренции с техникой по приемнику, ночное зрение проигрывает из-за отсутствия в нем кнопки «включить». Несомненным преимуществом электронного зрения является возможность получения цветных изображений уже при яркостях чуть превышающих пороговые. Наверное, к преимуществам камеры следует отнести возможность осознанной и свободной обработки первичного изображения, в то время как первичное низкоуровневое изображение после ретины обрабатывается в значительной степени подсознательно по, возможно совершенному, но жёсткому алгоритму. Единственное, что сейчас может внушить оптимизм любителям натуральных средств — это нарастающая угроза использования мощных радиоэлектронных помех против, в частности, и оптико-электронных устройств. Для зрения в силу электро-фотохимического принципа его функционирования эти помехи пока безопасны.

Литература

Глезер В.Д., Подвигин Н.Ф. Световая чувствительность зрительной системы / Физиология сенсорных систем. Ч. 1. Физиология зрения. Под ред. Гершуни Г.В. М.: Наука, 1971. 211 с.

Тибилов А.С., Шелепин Ю.Е. Сравнение результатов психофизических и физиологических исследований порога зрительного восприятия // Сенсорные системы. 2013. Т. 27. № 3.С. 271–276.

Шелепин К.Ю., Пронин С.В., Шелепин Ю.Е. Распознавание фрагментированных изображений и возникновение “инсайта” // Оптический журнал. 2015. Т. 82. № 10. С. 72–80.

Ala-Laurila P., Donner K., and Koskelainen A. Thermal activation and photoactivation of visual pigments // Biophysical J. 2004. V. 86. June. P. 3653–3662.

Ala-Laurila P., Rieke F. Coincidence detection of single-photon responses in the inner retina at the sensitivity limit of vision // Current Biology. 2014. V. 24. № 24. P. 2888–2898.

Barlow H.B. Retinal noise and absolute threshold // JOSA. 1956. V. 46. № 8. P. 634–639.

Barlow H.B. Temporal and spatial summation in human vision at different background intensities // J. Physiol. I958. V. I41. P. 337–350.

Baylor D.A., Nunn B.J., and Schnapf J.L. The photocurrent, noise and spectral sensitivity of rods of the monkey macaca fascicularis // J. Physiol. 1984. V. 357. P. 575–607.

Bouman M.A. and Van der Velden H.A. The two-quanta explanation of the dependence of the threshold values and visual acuity on the visual angle and the time of observation // JOSA. 1947. V. 37. № 11. P. 908–919.

Bouman M.A. and Van der Velden H.A. The quanta explanation of vision and the brightness impression for various times of observation and visual angles // JOSA. 1948. V. 38. № 3. P. 231–236.

Denton E. J., Pirenne M. H. The absolute sensitivity and functional stability of the human eye // J. Physiol. I954. V. I23. P. 4I7–442.

DeVries S.H., Baylor D.A. Mosaic arrangement of ganglion cell receptive fields in rabbit retina // J. Neurophysiol. 1997. V. 78. P. 2048–2060.

Graham C. H., Margaria R. Area and the intensity-time relation in the peripheral retina // Am. J. Physiol. 1935. V. 113. P. 299.

Hecht S., Shlaer S., Pirenne M. H. Energy, quanta, and vision // J. Gen. Physiol. 1942. V. 25. P. 819–840.

Herrick R. M. Increment thresholds for multiple identical flashes in the peripheral retina // JOSA.1973 V. 63. № 10. P. 1261–1265.

http://www.nuvucameras.com/

Jensen RJ. Involvement of glycinergic neurons in the diminished surround activity of ganglion cells in the dark-adapted rabbit retina // Vis Neurosci. 1991;6:43–53.

Okawa H. and Sampath A. P. Optimization of single-photon response transmission at the rod-to-rod bipolar synapse // Physiology. 2007. V. 22. August. P. 279–286.

Osterberg G. Topography of the layer of rods and cones in the human retina // Acta Ophthal Suppl. 1935. V. 6. P. 1–103.

Peichl L, Wässle H. The structural correlate of the receptive field centre of alpha ganglion cells in the cat retina //J Physiol. 1983;341:309–324.

Sakitt B. Counting every quantum // J. Physiol. 1972. V. 223. P. 131–50.

Sterling P., Freed M.A., and Smith R. G. Architecture of rod and cone circuits to the on-beta ganglion cell // J. Neuroscience. 1988. V. 8. № 2. P. 823–842.

Vos Johannes J. On the cause of disability glare and its dependence on glare angle, age and ocular pigmentation // Clin. Exp. Optom. 2003. V. 86. № 6. P. 363–370.

Глава 3.
Острота зрения и Краудинг-эффект у взрослых и детей различного возраста

В. М. Бондарко

Глава посвящена изучению свойств функциональных элементов, участвующих в опознании зрительных изображений на пределе разрешения зрительной системы, и исследованию взаимодействий между этими элементами. Будет рассмотрена нейронная сеть, обеспечивающая одну из важнейших функций зрения — видение объектов малых размеров. Выбор таких объектов не случаен, поскольку их опознание обеспечивают нейроны с рецептивными полями небольших размеров. Тем самым возникают естественные ограничения, упрощающие рассматриваемую нейронную сеть. В связи с такой постановкой задачи, основной акцент сделан на изучение зависимости оценок остроты зрения от выбора тестовых объектов и влияния на опознание объектов окружающих изображений (феномен краудинг-эффекта).

1. Разрешающая способность зрительной системы

Для оценки функционального состояния зрительной системы человека одним из основных показателей служит острота зрения. В офтальмологической практике остроту зрения у взрослых людей измеряют при помощи таблиц, составленных из оптотипов Снеллена, Ландольта или Головина-Сивцева. Оптотипы из этих таблиц представляют собой буквы и пациента просят назвать букву, т. е. опознать ее форму. Однако для того, чтобы назвать букву, надо иметь возможность увидеть мелкие детали объекта, такие как отдельные штрихи. Поэтому остроту зрения наблюдателя часто определяют на основании критерия Рэлея как разрешающую силу оптики глаза. Разрешающая способность в оптике характеризует возможность видеть раздельно изображения двух близко расположенных точек. В зрении исследование разрешающей способности включает несколько уровней зрительной системы. Двухточечный тест позволяет изучить разреша-

ющую способность системы сетчатка—зрительная кора. Для тестовых решеток более важным оказывается уровень зрительной коры, поскольку накопление сигнала происходит уже в пределах вытянутых рецептивных полей нейронов зрительной коры. Варьируя характеристики стимула и меняя задачу для наблюдателя, можно в психофизических исследованиях определять функциональную способность различных уровней зрительной системы. Открытые Хьюбелем и Визелем [Hubel, Wiesel, 1962] удлиненные рецептивные поля (РП) нейронов стриарной коры, так называемые детекторы полос и краев, накапливают сигнал вдоль своей возбудительной части и тем самым оптимизируют отношение сигнал/шум. Предполагается, что при наблюдении изображений натуральных сцен для выделения сигнала из шума, работают именно такие поля [Field, 1987]. Они могут играть и основную роль при опознании изображений малых размеров, с помощью которых определяют остроту зрения.

Предположение о том, что двухточечный тест позволяет изучить разрешающую способность вплоть до зрительной коры, было проверено следующим образом. Были сопоставлены данные по восприятию двухточечного теста и по определению минимальной длины линии, при которой происходит различение ее ориентации. Измерить ориентацию изображений зрительная система может с помощью вытянутых рецептивных полей, находящихся на уровне не ниже стриарной коры. В рецептивных полях нейронов первичной зрительной коры происходит суммация слабого сигнала от нескольких рецепторов с тем, чтобы выделить его из шума, но в отличие от рецептивных полей более низких уровней (ганглиозных клеток сетчатки и РП наружного коленчатого тела, имеющих круглую форму) на этом уровне суммация происходит в одной ориентации. Для выявления характеристик РП нейронов стриарной коры (самых высокочастотных детекторов полос) в психофизических экспериментах определяли минимальную длину линий, при которой наблюдатель мог отличить горизонтальную линию от вертикальной. Использовали линии с шириной 1.0 угл.мин, их длины менялась в пределах от 1.175 до 2–3 угл.мин с шагом 0.175 угл.мин. Стимул предъявляли бинокулярно в течение 500 мс на экране монитора с расстояния 8.2 м. Яркость экрана монитора была равна 80 кд/м^2, стимула—5 кд/м^2. Использо-

вали методику вынужденного выбора. Наблюдатель должен быть ответить, какую линию он видит: горизонтальную или вертикальную. Определяли вероятность правильного различения ориентации линий в зависимости от ее длины.

На рис. 1 показана зависимость процента правильного ответа от длины линии. Приведены данные четырех наблюдателей: АМ, НТ, ВЧ и ВБ. Видно постепенное увеличение процента правильного ответа от уровня случайного гадания до 85–97 % уровня правильных ответов. Выход на плато (этот уровень незначительно варьируется у разных наблюдателей) осуществляется при длине линии в пределах 1.75–2.25 угл.мин у разных наблюдателей. Можно полагать, что таким способом мы измерили размер возбудительной зоны самых высокочастотных детекторов полосы, имеющих РП наименьшего размера и располагающихся в стриарной коре. Полученные значения совпали с оценками восприятия двухточечного теста: при таких расстояниях между точками две точки видятся раздельно [Шелепин, Бондарко, 2002]. Результат позволил подтвердить предположение, что двухточечный тест позволяет проследить разрешающую способность зрительной системы, включая первичную зрительную кору. Данные по оценке минимальной длины линий, скорее всего, характеризуют уровень функционирования стриарной коры, однако, безусловно, здесь приведена упрощенная трактовка процессов, происходящих в зрительной системе. Далее проведено сравнение полученных минимальных длин линий с другими известными характеристиками.

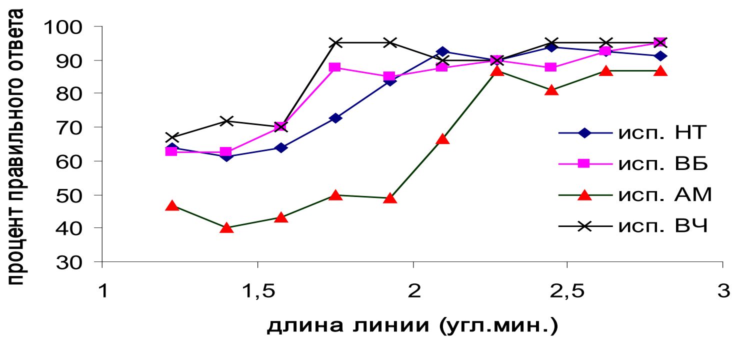


Рис. 1. Зависимость различения ориентации линий от длины

В настоящее время широкое распространение получила гипотеза, согласно которой первичное описание изображений осуществляется набором пространственно—частотных каналов, имеющих определенную ориентационную настройку [Campbell, Robson, 1968; Blakemore, Campbell, 1969]. При этом предполагается, что каналы образованы пространственными элементами. Термин “пространственный элемент” близок по своей сути понятию рецептивного поля. Разница заключается в том, что рецептивные поля определяют электрофизиологически—при отведении реакций отдельных нейронов. Пространственные элементы определяют в психофизических исследованиях—это означает, что они отражают работу всей зрительной системы. Ответ пространственных элементов на предъявляемый стимул равен интегралу от произведения функции, задающей контраст стимула, и весовой функции пространственного элемента. Нейрофизиологической основой пространственно-частотных каналов принято считать рецептивные поля нейронов стриарной коры (17 цитоархитектонического поля) или, по современной классификации, зоны V1. По своим свойствам пространственные элементы, образующие каналы, также близки к рецептивным полям стриарной коры.

Попытаемся объяснить полученные данные по различению ориентации функционированием наиболее высокочастотных пространственных элементов. Для этого рассчитаем их весовые функции. Воспользуемся данными, которые в более удобной аналитической форме представлены в работе Вилсона и Гелба [Wilson, Gelb, 1984]. На рис. 2, внизу, показано сечение двумерной весовой функции четного элемента—детектора полосы—в ориентации, перпендикулярной к оптимальной. В оптимальной ориентации данная функция принимает постоянные значения.

Как видно из рисунка, ширина возбудительной зоны пространственного элемента, расположенная в центре, составляет величину около 1.9 угл.мин, что согласовано с длиной линии, при которой наши испытуемые могут практически без ошибок определить ее ориентацию. Поэтому можно предположить, что в задаче различения ориентации линий малых угловых размеров принимают участие подобные элементы во взаимно перпендикулярных оптимальных ориентациях. При попадании линии в тормозные зоны элемент

с перпендикулярной оптимальной ориентацией даст больший ответ, поскольку вся линия будет находиться у него в возбудительной зоне.

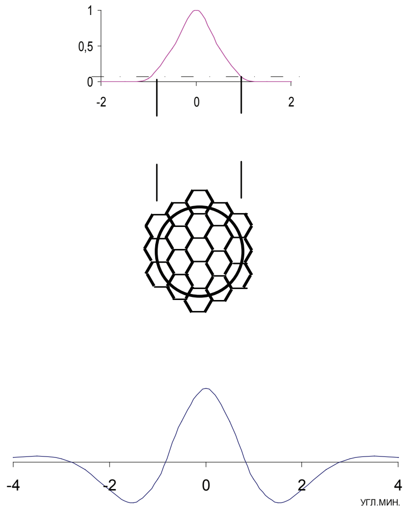


Рис. 2. Согласованность оптической функции рассеяния с мозаикой фоторецепторов и весовой функцией самых высокочастотных пространственных элементов

Сопоставим элемент дискретизации изображений, предложенный Ф. Кэмпбеллом и Ю. Е. Шелепиным [Campbell, Shelepin, 1989, 1992], с весовой функцией самого высокочастотного элемента. Отметим, что, несмотря на кажущуюся простоту и очевидность такого рассмотрения, до сих пор это никем сделано не было. На рис. 2 продемонстрированы результаты этого сопоставления. В верхней и средней частях показаны рисунки, аналогичные приведенным в работе Кэмпбелла и Шелепина. Здесь схематично изображена функция рассеяния и гексагональное расположение фоторецепторов на сетчатке. Функция рассеяния была измерена Кемпбеллом и Губишем [Campbell, Gubisch, 1966]. Ее эффективная часть на уровне 95 % охватывает область около 2 угл.мин, что соответствует размеру гексагона, изображенному ниже, в середине рисунка. Размер возбудительной зоны самого высокочастотного пространственного элемента, показанного на рис. 2 внизу, также равен этой величине.

Таким образом, весовая функция самого высокочастотного канала естественным образом ограничена физическими и физиологическими характеристиками зрительной системы: согласована с гексагональным расположением фоторецепторов на сетчатке и с оптической функцией рассеяния. Судя по этим результатам, согласованность должна проявляться и в соответствии размеров возбудительных зон самых мелких рецептивных полей ганглиозных клеток сетчатки и элемента дискретизации, поскольку подобные рецептивные поля фовеальной зоны сетчатки принимают участие в описании изображений любых форм на пределе разрешения.

Отметим, что представление о разрешающей способности во многом базируется на представлениях о роли дифракции при создании оптического изображения. В отличие от аберраций дифракцию преодолеть невозможно, так как она определяется фундаментальными волновыми свойствами света. Уменьшение дифракции возможно лишь за счет увеличения зрачка. Но большой зрачок приводит к увеличению искажений, создаваемых аберрациями. Теперь появилась техническая возможность уменьшить влияние аберраций. Но сможем ли мы при этом улучшить восприятие?

Здесь показано, что для нормальной остроты зрения требуется согласованное функционирование оптики, рецепторного и нервного отделов вплоть до стриарной коры. Нелишне напомнить, что зрительная система формируется в онтогенезе и в сформированном виде представляет собой структуру с согласованными связями. Какая произойдет перестройка всей этой системы связей при изменении какого-либо звена в общей цепи у взрослого человека? Во многом это неизвестно. Но для ряда случаев можно предсказать последствия. Например, при решении конкретной проблемы оптимизации зрения путем исправления аберраций, у некоторых пациентов может возрасти шум дискретизации изображения и при наблюдении мелких структур появится муар. Это возможно в том случае, если ширина функции рассеяния будет сужена до размера рецептора, вернее, до шага упаковки рецепторов. Во избежание этого, при дальнейшем развитии рефракционной хирургии, необходимо учитывать принципы согласования оптических и нейронных структур.

Острота зрения считается нормальной, если наблюдатель видит разрыв в кольце Ландольта, размер которого не превосходит 1 угловую минуту [Кравков, 1950]. На сетчатке одна угловая минута превышает размер одной колбочки, величина которой в центре фовеолы в среднем составляет 0.4–0.6 угловых минут. Таким образом, фовеолярная острота зрения эмметропического глаза лимитируется скорее дифракцией, а не размером колбочки, поскольку в зрительной системе эффективная ширина функции рассеяния (на половине высоты) перекрывает 2–3 колбочки. Подробные данные о строении сетчатки приведены в монографии Поляка [Polyak, 1941]. В ней показано, что в центральной части фовеолы размер колбочек минимален и равен 24 угловым секундам или около 1 микрона. В фовеоле вдоль горизонтальной оси может располагаться до 50 таких колбочек. При удалении от центра поля зрения размер колбочек и расстояние между ними увеличиваются. Такая структура сетчатки обеспечивает изменение остроты зрения человека при удалении от центра поля зрения.

2. Измерение остроты зрения с использованием различных оптотипов

В предыдущем разделе была рассмотрена разрешающая способность зрительной системы, которая во многом определяется оптикой глаза. Кроме оптических факторов на определение остроты зрения влияют также и психологические факторы, а именно, различие задач, поставленных перед пациентом или наблюдателем. Еще в XVIII веке Майер [Mayer, 1755/1987] показал, что при использовании разных тестов для измерения остроты зрения можно получить расхождение в два раза при измерении минимального угла разрешения. Расхождение между разными тестами известно среди офтальмологов и не раз отмечалось разными авторами. В своем обзоре Риггс [Riggs, 1965] рассматривает несколько типов задач, используемых при измерении остроты зрения. В этих задачах используются разные объекты, что и приводит к разным результатам. Автор рассматривает задачи на обнаружение, разрешение, локализацию и опознание.

В задачах на обнаружение измеряется минимальный размер объекта, при котором этот объект обнаруживается. Эти тесты обычно

состоят из точек или тонких линий. Получаемый результат сильно зависит от полярности контраста теста (темный объект на светлом фоне или наоборот) и от уровня освещенности. Хехт и Минц [Hecht, Mintz, 1939; цитировано в обзоре Риггса] показали, что в оптимальных условиях можно обнаружить присутствие темной линии на светлом фоне, если ширина этой линии всего 1/2 угловой секунды. Этот результат кажется естественным, поскольку показано, что обнаружение сигналов осуществляется наиболее чувствительными элементами зрительной системы, являющимися довольно-таки низкочастотными. Подобные элементы начинают функционировать, как только энергия стимула начинает превосходить пороговую величину, то есть яркий стимул (например, грозовая молния или звезда) с очень малой шириной линии или размером может вызывать реакцию.

О задачах локализации говорят тогда, когда от наблюдателя требуется определить наличие и направление смещения одной части объекта относительно другой части объекта. Измерения с помощью стимула Верньера дают величину остроты в пределах нескольких угловых секунд (около 3–6 угл.сек.), что намного меньше, чем расстояние между колбочками в центре поля зрения. Поэтому пороги различения, имеющие величину ниже 30 угл.сек, Вестгеймер [Westheimer, 1980] предложил назвать порогами “гипероcтроты”.

Задачи на опознание фактически являются единственными задачами, где наблюдатели должны видеть расстояние между объектами или его элементами. Типичными объектами для измерения этого типа остроты зрения являются две точки, две линии, решетки, образованные параллельными полосами, или шахматные паттерны. Выполнение этого типа задач также зависит от полярности контраста и уровня освещенности. Приводимые в обзоре Риггса [Riggs, 1965] данные варьируют от 30 угловых секунд до 3 угловых минут.

Еще в прошлом веке было показано, что при одинаковом размере буквы, имеющие больше линий, на пределе разрешения зрительной системы опознаются хуже. Поэтому для определения разрешающей способности Ландольт и Снеллен [Landolt, 1889; Snellen, 1862] стали использовать в качестве оптотипов одинаковые буквы, отличающиеся только ориентацией. Например, буквы Е в четырех ориента-

циях составили одну из разновидностей таблицы Снеллена, а кольца с разрывами — таблицу Ландольта. Несмотря на предпринятые усилия по совершенствованию таблиц, для офтальмологов не является новым тот факт, что оценка остроты зрения одного и того же пациента зависит от используемых тестов [например, Mayer, 1755/1987; Riggs, 1965].

В 1937 году Шлаер [Shlaer, 1937] показал, что в зависимости от освещения для людей с нормальным зрением оценка остроты зрения, полученная с помощью колец Ландольта, может быть ниже, чем острота зрения, измеренная с помощью прямоугольных решеток; они могут быть равны; или кольцо Ландольта может различаться лучше, если выражать остроту зрения в единицах разрыва кольца или расстояния между полосами решетки. Так, например, при освещенности около 300 троландов острота зрения, измеренная с помощь решеток, в 1.5 раза ниже, чем острота, измеренная с помощью колец Ландольта.

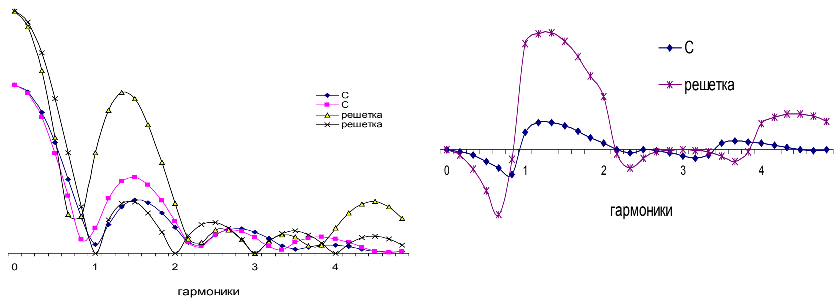
Первоначально аналогичные результаты были получены в исследованиях [Бондарко, Данилова, 1996] при изучении краудинг-эффекта. При измерении остроты зрения с помощью колец Ландольта оказалось, что один из тренированных наблюдателей правильно определял ориентацию разрыва при диаметре кольца в 2.25–2.7 угл.мин., то есть ширине разрыва в 0.45–0.54 угл.мин. В терминах пространственных частот такая величина разрыва соответствует частотам 66.7–55.6 цикл/град. Но ранее на этом же наблюдателе при использовании синусоидальных решеток было получено, что предельно видимая им частота составляет всего лишь 30–33 цикл/град. Другие наблюдатели, хотя и показали более низкие пространственные частоты, соответствующие разрыву кольца, но у всех эти частоты были выше 40 цикл/град.

Такие результаты вызвали естественные сомнения, не обусловлены ли различия в оценках несовершенством в подборе единиц измерения остроты. Это предположение было подвергнуто экспериментальной проверке и анализу.

Пространственно-частотный анализ оптотипов. Для устранения возникших расхождений в оценках остроты зрения был проведен Фурье-анализ зрительных стимулов, использовавшихся для опреде-

ления остроты зрения. Были вычислены амплитудные двумерные спектры изображений [Bondarko, Danilova, 1997; Бондарко, Данилова, 1998] прямоугольных решеток, имеющих две полосы, колец Ландольта и букв Е. Анализ спектров колец Ландольта и Снелленовских Е показал, что достаточно рассматривать только сечения спектров во взаимно перпендикулярных ориентациях. На рис. 3, А показаны сечения двумерных спектров кольца Ландольта и прямоугольной решетки в этих ориентациях. Здесь стимулы имеют одинаковый размер. По оси абсцисс отложены гармоники. Первая гармоника соответствует частоте, период которой равен размеру стимула, у второй гармоники период равен половине размера стимула и т. д. По оси ординат — амплитуда спектра, дана в линейном масштабе. Из рисунка видно, что первый локальный максимум у спектров решетки и у кольца Ландольта приходится на частоты, близкие 1.3–1.5 гармоникам. На рис. 3, Б приведены, кроме того, разницы между сечениями спектров в интересующих нас ориентациях. Видно, что для кольца Ландольта наибольшая разница в спектрах во взаимно перпендикулярных ориентациях наблюдается при частотах, с периодом в 1.15–1.3 раза превосходящих размер кольца. На тех же частотах наблюдается и разница в сечениях спектров для решетки. Разницы в спектрах этих стимулов близки по форме, они отличаются только по амплитуде. Необходимо заметить, что разрывы в кольце и решетке составляют полпериода частот, которые принято традиционно брать за единицы измерения. Поскольку в кольце Ландольта разрыв равен 1/5 его размера, то частота разрыва соответствует 2.5 гармонике, а частота разрыва решетки — 1.5 гармонике. Как видно из графиков на рис. 3, А максимальная амплитуда у спектров этих стимулов не соответствует этим гармоникам. Для кольца Ландольта максимум находится примерно на 1.4 гармонике, а для решетки — на 1.3 гармонике, примерно на тех же частотах наблюдается и максимальная разница в спектрах во взаимно перпендикулярных ориентациях (рис. 3, Б). Так как оптическая функция рассеяния воздействует на стимулы таким образом, что более высокие частоты в изображениях ослабляются, то на самом деле амплитуда на гармониках 2.5 и 1.5 и разница между ними еще меньше. Исходя из рассмотрения спектров, мы сделали вывод, что частоты разрывов (2.5 и 1.5 гармоники) не могут

быть использованы зрительной системой для определения ориентации этих стимулов, поскольку на этих частотах очень незначительная амплитуда в спектре и разница между амплитудами.



А Б

Рис. 3. Сопоставление амплитудных спектров кольца Ландольта и прямоугольной решетки

А — сечения двумерных амплитудных спектров прямоугольной решетки (решетка) и кольца Ландольта (С) во взаимно перпендикулярных ориентациях. *Б* — разница в спектрах для решетки и кольца Ландольта в перпендикулярных ориентациях

Экспериментальная проверка гипотезы о единицах измерения остроты. Мы предположили, что зрительная система для различения ориентации стимулов может использовать частоты, на которых наблюдается максимум у спектров. (Заметим, что такое предположение имеет основание, если придерживаться концепции о частотном анализе изображений, осуществляемом зрительной системой [Campbell, Robson, 1968; Blakemore, Campbell, 1969].) Для проверки предположения мы пересчитали полученные нами данные по определению минимально видимых размеров колец Ландольта и решеток с поправкой на их спектры в терминах минимально видимых частот. Минимально видимые размеры стимулов были получены в экспериментах как 75 % пороги правильных ответов при определении ориентации стимулов. Пороги были посчитаны как взвешенные средние от тех размеров, при которых процент правильных ответов на изолированный стимул был ниже и выше 75 %. В Таблице 1 для каждого испытуемого приведены предельные частоты, полученные на решетках и кольцах Ландольта двумя способами. В первых двух строках

таблицы приведены величины, соответствующие частотам разрывов в кольцах Ландольта и решетках — 2.5 и 1.5 гармоникам. В последующих строках эти величины пересчитаны с учетом поправки на спектры этих изображений. У кольца Ландольта предельной берется более низкая частота — 1.4 гармоники. У решеток тоже вводится поправка, так как максимум амплитуды в спектре решетки сдвинут в более низкочастотную область — на 1.3 гармоники (см. рис. 3). Из таблицы следует, что предельно видимые частоты с учетом поправок согласованы для этих оптотипов.

Таблица 1

Максимально видимые пространственные частоты, полученные путем пересчетов из размеров минимально видимых оптотипов

Испытуемый	ИК	МД	ВБ
Частота разрыва в решетке (цикл./град.)	28	33	39
Частота разрыва в кольце Ландольта	44	48	60
Частота решетки с учетом поправки	24	28	33
Частота кольца Ландольта с учетом поправки	24	27	33

Следует отметить, что в данных расчетах мы не учитывали оптическую функцию рассеяния, вид которой зависит от освещенности, а рассматривали только спектры изображений. При меньшей же освещенности частотная характеристика становится более низкочастотной. Также не учитывался размер зрачка. Видимо, рассматривая эти характеристики, можно получить и разницу в оценках измерений остроты при низких освещенностях, приведенных Риггсом [Riggs, 1965].

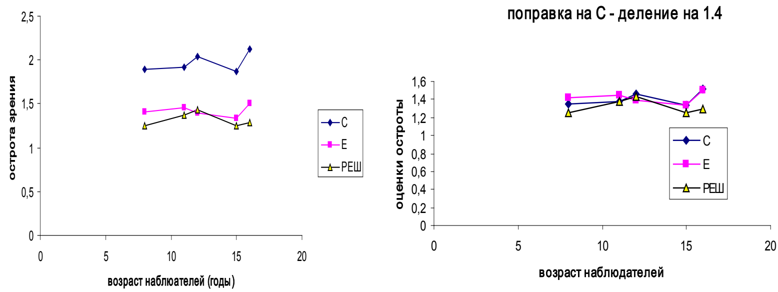
Проведенный анализ показывает, что не всегда острота зрения определяется разрешением двух точек или линий в смысле критерия Рэлея. Опознание формы изображения, как это происходит в случае определения ориентации кольца Ландольта, не требует четкого видения разрыва. Приведенный Морганом [Morgan, 1991] пример из астрономии подтверждает этот вывод: два объекта не разрешены в точном смысле критерия Рэлея, однако очень легко сделать вывод, что наблюдается два астрономических объекта, а не один, основываясь на форме суммарного распределении освещенностей двух объектов.

Эти исследования инициировали ряд последующих работ, которые подтвердили наши предположения об анализируемой зрительной системой частоте в данных оптотипах. В исследовании Хесса и др. [Hess et al., 2000] была произведена фильтрация изображений кольца Ландольта узкополосными частотными фильтрами шириной в 1 октаву. Данные по опознанию ориентации кольца, полученные на отфильтрованных изображениях, сопоставлялись с данными по опознанию неотфильтрованного кольца Ландольта на пределе разрешения зрительной системы наблюдателей. Было показано хорошее соответствие (процента правильных ответов) опознания неотфильтрованного изображения только тем отфильтрованным, в которых присутствовали частоты, близкие 1.15–1.3 гармоникам. При оставлении в изображении других частот (например, соответствующих разрыву) процент правильных ответов был близок к случайному уровню.

Анализ двух тестовых изображений — колец Ландольта и прямоугольных решеток — показал, что путем подбора соответствующих единиц измерения можно получить соответствие между оценками остроты зрения при использовании этих тестов. Для полноты картины мы также рассмотрели спектры букв Е из таблицы Снеллена [Bondarko, Danilova, 1997]. Оказалось, что частота разрыва (2.5 гармоника) и максимум в спектре у этого изображения примерно соответствуют друг другу. Таким образом, изолированные буквы Е дают более точную оценку остроты зрения, чем кольца Ландольта и решетки, если эти оценки выражать в общепринятых терминах как величину минимально видимых разрывов. Можно было бы ожидать, что острота зрения, измеренная с помощью таблицы Снеллена, составленной из букв Е, с единицей измерения, соответствующей величине разрыва, может быть более точной.

В дальнейшем были проведены измерения остроты зрения у большого количества школьников (около 300 человек). На рис. 4 приведены средние значения оценок остроты зрения с использованием в качестве оптотипов колец Ландольта, букв Е Снеллена и прямоугольных решеток [Бондарко, Семенов, 2005, 2012]. Данные по решеткам и буквам Е практически совпадают. Средняя острота зрения по ним имеет значение около 1.4 ед. Оценки остроты, опреде-

ленные с помощью колец Ландольта значительно и достоверно выше (рис.4, А). Они практически совпадают с оценками, полученными при использовании решеток и букв Е, если поделить их значения на 1.4 (рис. 4, Б). Необходимо отметить, что эти данные были получены при предъявлении изображений на дисплее с яркостью экрана 80 кд/м^2, стимулов –5 кд/м^2. При измерении остроты зрения с помощью отпечатанных на лазерном принтере таблиц с шагом между строками в 0.1 ед., любезно предоставленных нам Г. И. Рожковой, разница в оценках остроты зрения при применении колец Ландольта и букв Е Снеллена оказалась также достоверно большей для колец, чем для букв Е, но составила меньшую величину [в 1.1. раз, Бондарко, Семенов, 2005] при сопоставимых значениях яркости. Возможно, что сказалось рассеивание света на не совсем гладкой бумаге или наличие бликов от приборов освещения. Методика измерения остроты в обоих случаях была одинаковой.



А Б

Рис. 4. Средние оценки остроты зрения у школьников, полученные с использованием разных оптотипов.

А — оценки остроты, определенные по величине минимально видимого разрыва в тестовых изображениях при использовании тестовых колец Ландольта (С), знаков Снеллена (Е) и прямоугольных решеток (реш). *Б* — проведена корректировка оценок остроты для тестовых колец Ландольта с учетом спектра этих изображений

В настоящем разделе не затрагивается вопрос об изменении остроты зрения у детей с увеличением возраста. На приведенных графиках острота зрения одинакова у школьников разных классов. Однако, по нашим данным, результаты могут существенно зависеть

от выборки детей [Бондарко, Семенов, 2005]. В этой работе сравниваются данные по измерению остроты зрения у школьников поселка Колтуши и г. Санкт-Петербурга. Они достоверно различаются: в сельской школе процент детей с ухудшенным зрением не зависит от возраста, в то время, как у городских школьников детей с ухудшенным зрением становится больше в старших классах.

Неоднозначные данные по измерению остроты зрения с помощью различных оптотипов, а также влияние пространственно-частотного состава изображения на измерение остроты зрения привело к попыткам создания новых оптотипов, учитывающих недостатки хорошо известных тестов. Подобными оптотипами могут являться различные решетки или отфильтрованные изображения букв, в которых сохранены высокие частоты и подавлены низкие и средние. Приближаясь или удаляясь от теста, наблюдатель отмечает внезапное видение (опознание) этого оптотипа лишь с определенного расстояния, зависящего от остроты зрения данного наблюдателя. Наблюдатель может узнать изображение, пользуясь исключительно механизмами высоких пространственных частот. Догадаться о форме изображения по низкочастотным составляющим, присутствующим в буквах обычного типа, в этой ситуации невозможно, так как именно низкочастотные составляющие отсутствуют в синтезированных оптотипах [Шелепин и др., 1987].

Несмотря на большую точность, получаемую с помощью новых оптотипов, они не получили достаточного распространения, так как в клинической практике офтальмологу важно иметь возможность быстро определить наличие нарушений в зрительной системе, что возможно и с помощью старых методов. Новые оптотипы дают возможность более детально исследовать зрительную систему. Однако при изготовлении новых таблиц возникают большие трудности. Здесь следует также отметить, что измерение остроты зрения с помощью любых тестов дает информацию о состоянии только, в основном, одного участка зрительной системы человека — оптики глаза и частично сетчатки. Большую диагностическую ценность представляет измерение контрастной чувствительности [Волков и др., 1983; Шамшинова, Волков, 1998; Шелепин и др., 1985], а не только остроты, которая соответствует только измерению одной точки на

кривой контрастной чувствительности. Простейший вид таблиц для измерения контрастной чувствительности — таблицы Пелли- Робсона, состоящие из букв латинского алфавита различного контраста. Для измерения остроты зрения у детей, не знающих буквы или не владеющих речью, используются другие символьные тесты [см. монографию Хювяринен, 1996].

3. Краудинг-эффект и острота зрения

Как было сказано в предыдущем параграфе, на определение остроты зрения наблюдателя даже с нормальной оптикой влияют различные параметры, среди которых наибольшее влияние оказывают вид теста и условия освещенности. Однако существует еще один фактор, определяющий нашу способность опознавать форму на пределе разрешения зрительной системы. Это влияние рядом расположенных объектов. В ежедневных ситуациях появление какого-либо одиночного символа (буквы или цифры) — это скорее исключение, чем правило: буквы обычно группируются в слова, а цифры — в группы цифр. С прошлого века известно, что восприятие групп, составленных из букв или цифр, нельзя предсказать, основываясь только на восприятии отдельных символов. В тридцатых годах появился цикл из пяти работ Вудроу, посвященный изучению восприятия стимулов, состоящих из двух или четырех букв, различным образом расположенных в зрительном поле и предъявляемых на 0.1 сек. Изменялось расстояние между буквами, их взаимное расположение и их группировка. Была обнаружена достоверная разница в ответах на 2-х и 4-х буквенные слова, и автор предположил, что эту разницу можно объяснить конфигурацией самого паттерна. В последней из 5 работ [Woodrow, 1938] был сформулирован основной вывод: этот эффект зависит от 1) распределения внимания, 2) смешиваемости или ‘торможения узнавания’, являющегося результатом пространственной близости букв и 3) степени видимости букв.

Этот вывод в общих чертах формулирует основные предположения, появившиеся гораздо позднее для объяснения явления, получившего название краудинг-эффект [Ehlers, 1953; Stuart, Burian, 1962]. Само явление заключается в ухудшении опознания тестово-

го символа, окруженного буквами или цифрами. При изучении краудинг-эффекта часто для обозначения дополнительных объектов, отвлекающих наше внимание от теста, применяют слово “дистрактор”. Этот термин также будет использоваться в дальнейшем при изложении материала.

Краудинг-эффект имеет ряд пространственных характеристик. Впервые подробно эти характеристики были изучены Фломом и соавторами [Flom et al., 1963a, б]. В качестве стимулов эти исследователи использовали кольца Ландольта в 4-х ориентациях, окруженные на разном расстоянии полосами (рис. 5). Наблюдатель должен был определить ориентацию тестового кольца. Исследования проводились на пределе разрешения зрительной системы. В работах было показано следующее: 1) ухудшение восприятия наблюдается только при близких расстояниях между тестовым объектом и окружением, примерно равных 4/5 размера минимально видимого кольца, 2) при расстояниях, больших размера кольца, не наблюдается существенных отличий процента правильных ответов от величины, полученной при обнаружении ориентации одиночного тестового кольца, 3) маскировочная функция зависимости процента правильных ответов от расстояния до окружения имеет s-образную форму (при самых близких расстояниях между дистракторами и стимулом у многих испытуемых наблюдается улучшение опознания — увеличение процента правильных ответов по сравнению с максимальным ухудшением восприятия при более дальних расстояниях). Оказалось, что маскировочные кривые разных наблюдателей подобны; они накладываются друг на друга [Flom et al., 1963a], если их пронормировать к минимальному углу разрешения каждого испытуемого.

Флом и его коллеги [Flom et al.,1963 a, б] предположили, что взаимодействие контуров имеет нейронный базис. Был исследован вопрос, на каком уровне происходит это взаимодействие: на уровне сетчатки или где-нибудь выше в зрительной системе [Flom et al.,1963 б]. Для этого использовалась методика контралатерального предъявления: тестовые кольца Ландольта предъявлялись одному глазу, а окружающие контуры — другому глазу с помощью соответствующим образом расположенных поляроидных фильтров. В результате были получены кривые зависимости краудинг-эффекта при

контралатеральном предъявлении тестового стимула и окружающих его полос. Они практически совпали с кривыми, полученными при ипсилатеральном предъявлении. Сходство между ипси- и контралатеральными эффектами делает мало вероятным существование отдельных локусов взаимодействия на уровне сетчатки и свидетельствует о том, что это взаимодействие имеет место на том уровне зрительной системы, где соединяется сенсорная информация от обоих глаз, то есть не ниже стриарной коры.

Несмотря на то, что работ по изучению краудинг-эффекта в центре поля зрения и на пределе разрешения зрительной системы сравнительно мало, а данные, полученные в разных экспериментах, трудно сравнивать, так как использовались различные процедуры измерения и тестовые стимулы, некоторые общие закономерности все же можно сформулировать следующим образом:

1) влияние окружающих объектов уменьшается при удалении их от опознаваемого объекта [Flom et al., 1963 a, b; Бондарко, Данилова, 1996];
2) эффект имеет большую протяженность на периферии, чем в центре поля зрения, т.е. размер зоны взаимодействия на периферии значительно больше, чем в фовеа [Bouma, 1970; Nazir, 1992; Strasburger et al., 1991];
3) влияние дистракторов асимметрично при предъявлении стимула не в центре поля зрения: дополнительные контуры, расположенные с периферической стороны тестового стимула ухудшают распознавание слабее, чем расположенные ближе к фовеа [Bouma, 1970];
4) эффект также зависит от того, где предъявляется стимул: справа и слева в поле зрения краудинг-эффект слабее, чем на том же расстоянии сверху или снизу [Bouma, 1970];
5) объекты разной формы приводят к разному виду маскировочной функции [Бондарко, Данилова, 1996].

В недавних обзорах [Atkinson, 1991; Flom, 1991; Levi, 2008], посвященных краудинг-эффекту, авторы пишут о том, что, несмотря на длительный период исследований, нет удовлетворительного объяснения этому феномену. Леви и др. [Levi et al., 2002] предполагают, что в фовеа краудинг-эффект является простой маскировкой,

в отличие от периферии. В частности, тормозное влияние проявляется на расстояниях, пропорциональных размеру стимула. Несогласие с этой гипотезой высказывают ряд авторов, показывая, что даже в фовеа дело обстоит сложнее: у краудинг-эффекта проявляются различные свойства в зависимости от тестовых стимулов и окружения [Бондарко, Данилова, 1999; Данилова, Бондарко, 2002], их размеров и частотного состава [Chung et al., 2001], несопоставимые с обычной маскировкой.

Для исследования этого явления на пределе разрешения зрительной системы мы использовали кольца Ландольта, окруженные полосами той же ширины и длины, кольцами Ландольта той же величины или прямоугольными решетками, буквы Е Снеллена, окруженные аналогичными буквами и прямоугольные решетки. На рис. 5 приведены примеры этих стимулов.



Рис. 5. Примеры используемых для изучения краудинг-эффекта стимулов: кольцо Ландольта окружено полосами, буква Е окружена аналогичными буквами, решетки окружены решетками с той же или более высокой пространственной частотой. В экспериментах меняется расстояние между тестовым стимулом и окружением

На рис. 6 сравниваются кривые, полученные для центральных колец Ландольта и решеток, имеющих разное окружение. У испытуемого ИК приведена кривая при окружении кольца полосами, а у испытуемых ВБ и МД—аналогичными кольцами. Отдельные точки слева соответствуют проценту правильных ответов на изолированный стимул. На рисунке вертикальными линиями обозначены те минимальные расстояния до дистракторов, при которых ответы уже не отличаются достоверно от ответов на изолированный стимул (где кривые выходят на плато). Заметим, что для колец Ландольта уровень случайного гадания составляет 25 %, так как выбор был из

четырех ориентаций колец (верх, низ, лево, право), а для решеток — 50 %. Поэтому кривые сильно расходятся, особенно при близких расстояниях до окружения.

У всех испытуемых влияние окружающих изображений на определение ориентации решеток сказывается на большем протяжении, чем влияние окружения на опознание ориентации колец Ландольта. Приведенные на рисунке для данных тестовых частот размер тормозной зоны при окружении тестовых решеток решетками той же частоты у наблюдателя ВБ был равен 3.4, у МД — 3.8, у ИК — 3.2 угл.мин. Размер зоны при окружении тестовых решеток квадратами у испытуемого ВБ равен 2.6, у МД –2.9, у ИК — 2.1 угл.мин. Размер зоны при использовании тестовых колец Ландольта у наблюдателя ВБ равен 2.1 угл.мин (размер кольца — 2.7 угл.мин), у МД –2.4 угл.мин (кольцо — 3.15 угл.мин), у ИК — 3.4 угл.мин (кольцо — 3.4 угл.мин).

Были сделаны попытки объяснить расхождения в размерах зон, полученных при использовании в качестве тестовых стимулов колец Ландольта и решеток, а также разницу в форме маскировочных функций. Большее основание для обсуждения имеют два возможных предположения. Первое — наиболее простое — зона взаимодействия формируется под воздействием тестового стимула. То есть, функционируют рецептивные поля, согласованные с изображением. Таким образом, размер зоны взаимодействия будет зависеть от исходного стимула. На возможность согласованной фильтрации указывают некоторые психофизические исследования и модельные эксперименты [Hauske et al., 1976].

Второе предположение использует гипотезу о пространственно-частотном анализе изображений [Campbell, Robson, 1968], осуществляемом зрительной системой, и гипотезу о взаимодействии между пространственно-частотными каналами [Polat, Sagi, 1993]. Поскольку зоны при использовании тестовых решеток получились разными для окружающих их решеток той же самой частоты и более высокой, то это предположение также может иметь место для случая определения ориентации решеток. Напомним, что в рамках этой гипотезы торможение должно сказываться сильнее при окружении тестового стимула изображениями, имеющими тот же спектральный состав.

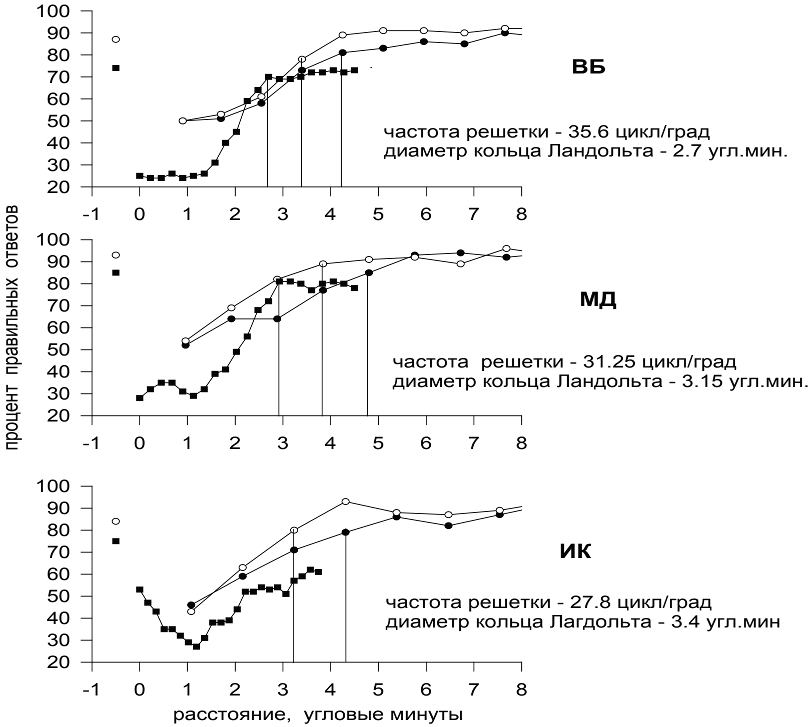


Рис. 6. Влияние окружения на определение ориентации тестовых стимулов

Тестовые стимулы — решетки (круги) или кольца Ландольта (квадраты). У наблюдателей ВБ и МД кольца Ландольта окружены аналогичными кольцами, у наблюдателя ИК — полосами. Сплошные круги — окружение аналогичными решетками, полые круги — высокочастотными решетками. Вертикальными линиями обозначены минимальные расстояния, при которых не было значимых отличий ответов на стимулы с окружением от ответов на изолированные стимулы. Отдельные точки слева — процент правильных ответов на изолированные стимулы

Чтобы объяснить полученные результаты с тестовыми кольцами Ландольта, мы рассмотрели функционирование самых высокочастотных элементов (или рецептивных полей — РП) зрительной системы. Для этого были посчитаны весовые функции (ВФ) высокочастотных элементов, используя приведенные в работе Вилсона

и Гелба [Wilson, Gelb, 1984] формулы и параметры для самого высокочастотного канала. Следуя модели Вилсона и Гелба, мы предположили, что в перпендикулярной ориентации ВФ пространственного элемента является постоянной функцией. Такое предположение о двумерной ВФ элемента оправдано нейрофизиологическими данными. Двумерные ВФ элементов «детекторов» полосы (четного элемента) и края (нечетного элемента) показаны на рис. 7.

На рисунке 8 показано схематично функционирование подобных элементов при предъявлении изолированного кольца Ландольта в разных ориентациях. Элемент имеет вертикальную оптимальную ориентацию и три зоны: в центре возбудительную и две тормозные по бокам. Причем ширина возбудительной зоны примерно равна ширине тормозной. Здесь размер кольца Ландольта взят равным размеру возбудительной зоны элемента.

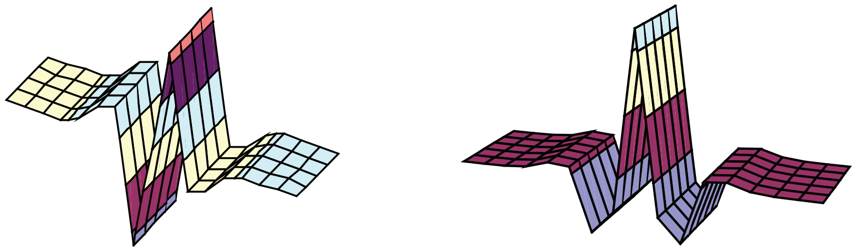


Рис. 7. Двумерные весовые функции высокочастотных «детекторов» края и полосы.

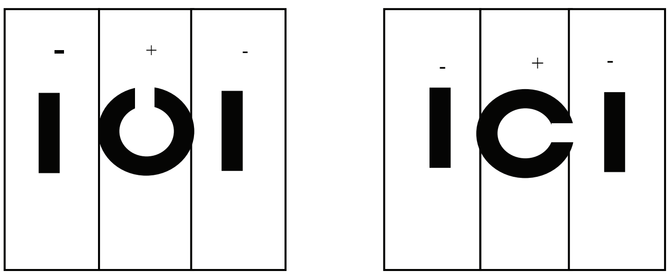


Рис. 8. Схема взаимодействия рецептивного поля с тестовым изображением кольцом Ландольта и дистракторами — полосами.

Ответ таких элементов будет меньше при предъявлении кольца с разрывом вверху (или внизу), так как разрыв приходится на максимум функции «детектора» полосы, а «детекторы» края не реагируют на данный стимул из-за его симметричности. При нахождении разрыва у кольца справа (или слева) ответ «детектора» полосы будет больше, поскольку разрыв попадает на тот участок ВФ, где ее значения малы. Кроме того, при таком положении кольца начинает давать ответ и «детектор» края, поскольку стимул асимметричен. Таким образом, возникает разница между ответами элементов при предъявлении кольца с разрывами «верх-низ» или «право-лево».

В рассмотренном примере элементы, имеющие оптимальную вертикальную ориентацию, дадут больший ответ на кольца, повернутые вправо и влево. Соответственно другие элементы, имеющие оптимальную горизонтальную ориентацию, дадут больший ответ на кольца, повернутые вверх и вниз. Таким образом, совместное функционирование элементов «детекторов» краев и полос во взаимно перпендикулярных ориентациях позволяет определить положения разрывов в кольце Ландольта «верх-низ» или «право-лево». Положение же разрыва «право» или «лево» (аналогично верх или низ) определяют «детекторы» краев одинаковой ориентации, имеющие ВФ противоположного знака.

Эта схема дает довольно хорошее объяснение краудинг-эффекту для тестового кольца Ландольта. В самом деле, по нашим данным и данным Флома [Flom et al., 1963a] размер тормозной зоны примерно соответствует размеру тестового кольца. У приведенных на рис. 7 элементов размеры возбудительных и тормозных зон примерно одинаковы. При попадании дистракторов в тормозную зону элементов (рис. 8) происходит уменьшение ответа на тестовое кольцо. При этом при симметричном предъявлении дистракторов относительно тестового кольца ответ «детектора» края равен нулю. Прямым подтверждением этой схемы являются полученные нами и другими исследователями маскировочные функции, отражающие зависимость правильных ответов от расстояния до различного окружения [Бондарко, Данилова, 2002]. В самом деле, использование узкой полосы в качестве стимула аналогично использованию δ-импульсов в технике. Флом и др. [Flom et al., 1963a, б] и мы [Бондарко, Данилова,

1996; Danilova, Bondarko, 2007] использовали окружение кольца Ландольта узкими полосами. Был получен характерный вид маскировочной функции, на котором видно сначала ухудшение опознания ориентации кольца, а затем улучшение (см. рис. 6, наблюдатель ИК). Такой вид маскировочной функции получается при функционировании высокочастотного четного элемента («детектора» полосы). Он в точности повторяет вид ВФ в тормозной зоне. (Заметим, что нечетные элементы не оказывают никакого воздействия на вид маскировочной функции, так как их ответы на дистракторы равны нулю при симметричном предъявлении дистракторов.) Для подтверждения этого утверждения мы посчитали маскировочные функции для всех типов окружения и сравнили их с экспериментальными [Danilova, Bondarko, 2007]. Расчетные и экспериментальные маскировочные функции соответствует друг другу. Таким образом, функционирование подобных элементов хорошо объясняет различный вид у маскировочных функций и одинаковый размер тормозных зон, полученных при применении разных дистракторов. Однако, полностью объяснить тормозный эффект функционированием линейных элементов на данный момент не представляется возможным.

Таким образом, краудинг-эффект с использованием в качестве тестовых стимулов колец Ландольта хорошо объясняется функционированием самых высокочастотных элементов. При этом размер этих элементов — «детекторов» полос — у каждого наблюдателя примерно соответствует трем размерам минимально видимого кольца. Однако не всегда функционирование подобных элементов объясняет ухудшение восприятия стимулов при близком расположении окружения. Механизм краудинг-эффекта для данного тестового стимула может быть обеспечен посредством функционирования самых высокочастотных пространственных элементов «детекторов» полос и краев, описывающих изображение, без введения добавочного латерального взаимодействия между ними. В рамках этих представлений ухудшение в опознании кольца Ландольта наблюдается при попадании дистракторов в тормозную зону элементов, описывающих тестовый стимул.

Краудинг-эффект для тестовых букв Е. Краудинг-эффекта был исследован также при использовании в качестве тестовых стиму-

лов Снелленовских букв Е. Напомним, что часто используемые таблицы, составленные из букв Е, в каждой строке имеют расстояние между буквами, равное удвоенной или утроенной ширине линий у букв (сейчас большое распространение получили и таблицы, в которых расстояния между символами равны их ширине). То есть, расстояния между буквами достаточно малы, и при оценке остроты зрения в таких условиях может сказываться краудинг-эффект. На рис. 9 показаны результаты измерений краудинг-эффекта у трех наивных наблюдателей ТА, ИР и ЕЛ. Изолированные точки слева — процент правильных ответов на стимул без окружения.

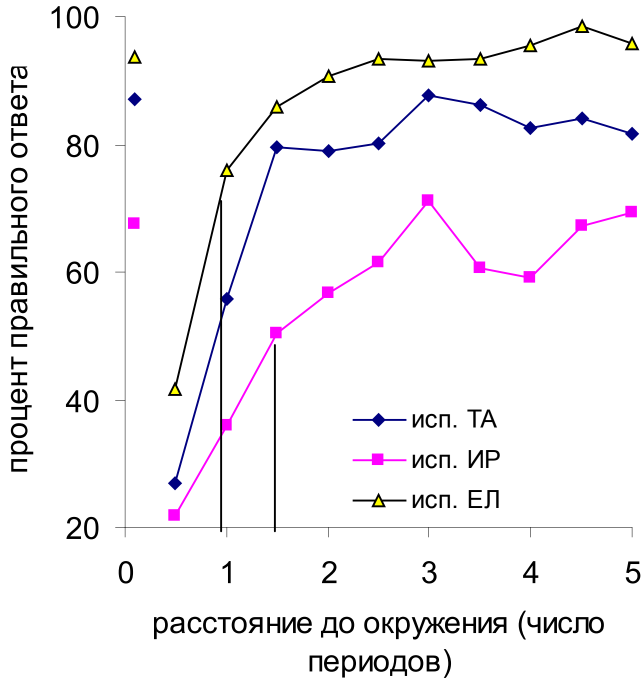


Рис. 9. Определение ориентации буквы Е в зависимости от расстояния до окружения

Соседние буквы оказывают воздействие на оценку ориентации тестового стимула при расстояниях, равных 1–1.5 периодам тесто-

вой частоты, или 2–3 ширинам линий. Полученные по абсолютному размеру или выраженные в периодах критической пространственной частоты зоны оказались сопоставимыми с зонами, определенными при изучении краудинг-эффекта с использованием тестовых колец Ландольта. Таким образом, использование изолированных букв Е и букв Е в таблице с малым расстоянием между буквами может приводить к разным оценкам остроты зрения у наблюдателей. Улучшение оценок остроты при тренировке может быть связано с уменьшением размеров тормозного взаимодействия между изображениями. В настоящее время в наших экспериментах у большой группы взрослых наблюдателей мы получили существенное сужение тормозных зон при использовании в качестве тестовых стимулов решеток. Таким образом, в задачах типа краудинг-эффекта проявляется пластичность зрительной системы и ее способность к обучению выполнения сложной задачи выделения тестового стимула из окружения на пределе разрешения зрительной системы.

Показана зависимость процента правильного ответа от расстояния до окружения, выраженного в количестве периодов оптимальной пространственной частоты. Период такой частоты равен ширине двух полос буквы Е.

Для тестовых решеток мы проанализировали гипотезу о влиянии на краудинг-эффект взаимодействия между частотными каналами. С нашей точки зрения, более оправданным является представление о том, что в ряде случаев при использовании тестовых решеток функционируют высокочастотные детекторы решеток, имеющие большие размеры, чем детекторы полос и краев. За счет чего, влияние окружения сказывается на больших расстояниях. Этот вывод основан был на данных, в которых для тестовых решеток менялась ориентация окружения. В одном случае ориентация окружающих решеток была фиксированной, аналогичной ориентации полос у колец Ландольта (см. рис. 5), в другом — ориентация окружающих решеток была случайной (вертикальной или горизонтальной), как в случае окружения букв Е. При фиксированной ориентации окружения ухудшение восприятия наблюдалось при тех же расстояниях, что и для колец Ландольта и букв Е, в то время, как для случайной ориентации эти расстояния оказались примерно в два раза больши-

ми [Данилова, Бондарко, 2002]. С точки зрения пространственно-частотного анализа и в том, и в другом случае окружение имело одинаковые характеристики (окружающие решетки имели ту же или более высокую частоту), то есть взаимодействие между каналами должно было проявиться в обоих случаях. Но этого не наблюдалось. Поэтому была высказана гипотеза, что при фиксированной ориентации окружения функционируют главным образом детекторы полос, а при случайной — детекторы решеток. Детекторы решеток могут быть более приемлемыми в последней ситуации, поскольку они не чувствительны к фазовой характеристике изображений. При случайном изменении ориентации у окружающих решеток в большинстве случаев изображения центральной решетки с окружением несимметричны, поэтому симметричные детекторы полос на них могут плохо реагировать, а ответы детекторов краев на симметричную центральную решетку той же ориентации равны нулю. Поэтому может возникнуть рассогласованность в восприятии ориентации центрального стимула.

Рассмотрев влияние различного окружения на разные тестовые стимулы на пределе разрешения зрительной системы, можно сделать следующие выводы, ранее не отмеченные другими исследователями:

1) Расстояние между тестовым стимулом и окружением, при котором наблюдается ухудшение восприятия, зависит от используемого тестового стимула, от спектрального состава окружения и расположения дистракторов.
2) Форма маскировочной функции при использовании в качестве тестовых стимулов колец Ландольта зависит от применяемых дистракторов.
3) Механизм краудинг-эффекта при использовании в качестве тестовых стимулов колец Ландольта можно описать с помощью функционирования самых высокочастотных элементов, присутствующих в зрительной системе данного наблюдателя, — «детекторов» полос и краев.
4) Тормозное воздействие окружения на определение ориентации букв Е распространяется на том же расстоянии, что и для колец Ландольта, равном 1–1.5 периодам тестовой частоты.

5) Наблюдаемое ухудшение восприятия тестовых решеток, окруженных близко расположенными аналогичными изображениями со случайной ориентацией, можно объяснить функционированием пространственных элементов — «детекторов» решеток или тормозным взаимодействием между самыми высокочастотными каналами.

4. Острота зрения и краудинг-эффект у детей различного возраста

Здесь представлены результаты измерения остроты зрения у детей двух возрастных групп различными тестами: у детей в возрасте от 4 до 18 месяцев и от 3 до 16 лет. У детей старшего возраста были изучены и характеристики краудинг-эффекта, что позволило выявить динамику изменения высокочастотных пространственных элементов в онтогенезе. Из-за сложности этой задачи мы не могли провести аналогичные исследования у детей младшего возраста. Однако мы приводим полученные другим методом данные у детей младшей возрастной группы ввиду того, что они также отражают изменения в зрительной системе и показывают важность измерения остроты зрения различными способами у детей самого раннего возраста, начиная от рождения. Кроме того, результаты этих исследований подтверждают положение о том, что при использовании разных тестов проявляется функционирование различных каналов зрительной системы.

Выбор именно таких возрастных групп детей не случаен. Хорошо известно, что наиболее интенсивно развитие зрения происходит в первые месяцы жизни ребенка. Хотя зрительная система малыша начинает функционировать с момента его появления на свет, только к 6-месячному возрасту зрение малыша его способность видеть окружающие предметы (различать в них мелкие и не очень контрастные детали, различать цвета, оценивать удаленность видимого объекта и т.д.) становится в какой-то мере подобной способности взрослого человека [Teller, 1997].

В первый год жизни острота зрения возрастает в 20–30 раз. Такой бурный рост этого показателя связан как с изменениями в периферических, так и центральных уровнях зрительной системы, по-

скольку острота зрения зависит от многих факторов: от рефракции, от развития аккомодации, от зрелости сетчатки, особенно — ее фовеальной области, от миелинизации нервных волокон, от зрелости корковых структур и т. д.

Поэтому изучение изменений в остроте зрения очень важно, начиная как можно с более раннего возраста. Интерес к этой проблеме обусловлен тем, что изучение физиологической функции в процессе ее становления и развития дает надежду лучше понять механизмы осуществления данной функции. Кроме того, эти исследования имеют и важное практическое значение: чем раньше обнаружены у ребенка те или иные дефекты развития, тем легче они могут быть выправлены.

Одной из основных причин очень низкой остроты зрения у новорожденных считается недоразвитость сетчатки. Сетчатка новорожденных сравнительно хорошо развита только на периферии. В центральной ее части отсутствует фовеальная ямка. Плотность колбочек в фовеальной области примерно в 3 раза меньше, чем у взрослого человека, а сами колбочки имеют короткие наружные членики, в которых содержится фотопигмент, — они примерно в 15 раз короче, чем у взрослого — [Yuodelis, Hendrickson, 1986]. Малое количество фотопигмента обуславливает низкую чувствительность зрения новорожденных, а редкое расположение колбочек в фовеальной области является одним из основных факторов, ограничивающим пространственное разрешение.

В первые месяцы жизни младенца изменения в сетчатке сопровождаются ростом в мозгу аксонов и дендритных отростков. Плотность синапсов в зрительной коре резко возрастает от рождения и до примерно 8-месячного возраста [Huttenlocher et al., 1982]. Период бурного нарастания синаптических связей в стриарной коре по времени совпадает с формированием у младенца ориентационной и пространственно частотной избирательности корковых рецептивных полей. В то же время меняется форма функции контрастной чувствительности: отмечается спад чувствительности для низких пространственных частот, у младенца появляется ряд функций, для осуществления которых необходимы тормозные взаимодействия (бинокулярное соперничество, стереопсис, дискриминация фазо-

вых соотношений, гипероострота для верньерных стимулов и.т.д.) Интересно отметить, что после восьмимесячного возраста начинает превалировать обратный процесс: постепенное уменьшение числа синапсов. Стабильный взрослый уровень их плотности в стриарной коре достигается примерно к 10-летнему возрасту [Huttenlocher et al., 1982].

Возраст обследованных нами детей охватывает верхние и нижние границы периода, в котором, судя по описанным выше фактам, могут существенно меняться оценки остроты зрения. К сожалению, выбор использованных тестов у детей младшей возрастной группы был ограничен тем, что для детей, не владеющих речью, невозможно использовать общепринятые тесты для определения остроты. Поэтому мы применяли разные тесты для детей этих возрастных групп. В обсуждении к разделу, посвященного оценке остроты зрения у детей младшего возраста, приводятся некоторые данные, полученные другими исследователями разными тестами.

4.1. Измерение остроты зрения у детей в возрасте от 4 до 18 месяцев

При измерении остроты зрения у детей малого возраста исследователи сталкиваются с рядом трудностей. Они обусловлены в первую очередь тем, что невозможно использовать традиционные психофизические методы для оценки остроты зрения: ребенку такого возраста нельзя объяснить задачу и получить от него словесный ответ. Поэтому измерения остроты зрения, контрастной чувствительности и других зрительных функций у малышей, не владеющих речью, базируются обычно на регистрации движений глаз (чаще всего — на предпочтительной фиксации), на регистрации вызванных потенциалов, оптокинетического нистагма или каких-либо вегетативных реакций.

Фантц [Fantz, 1961] показал, что малыши с момента рождения реагируют на изображения, где есть перепады освещенности и, в частности, — на решетки, предпочитая их гомогенному полю с такой же средней яркостью. В современных исследованиях пространственного зрения — для измерения функции контрастной чувствительности и остроты зрения обычно используют решетки с си-

нусоидальным или прямоугольным профилем распределения освещенности. Давидой Теллер разработаны тесты и процедура, позволяющая измерять эти показатели в поведенческих экспериментах с младенцами с высокой степенью точности и надежности [Teller et al., 1974; Teller, 1979; 1997]. Малышу предъявляют решетку и рядом с нею «бланк» — «пустой» стимул, уравненный с решеткой по средней яркости. Взрослый наблюдатель, который не видит, с какой стороны находится решетка, пытается определить это по поведению малыша — по направлению первого поворота глаз, по частоте фиксаций, общей их длительности, по эмоциональным реакциям ребенка. Вероятность правильных ответов наблюдателя в условиях такого вынужденного двухальтернативного выбора дает возможность применять обычные психометрические методы определения порогов — метод константных стимулов, метод лестницы.

При измерении остроты зрения по решеткам определяют максимальную пространственную частоту решетки (в циклах на градус), при которой еще удается установить ее предпочтение перед гомогенным полем. В офтальмологической практике кроме измерения остроты зрения по разрешению решеток используют и другие методы, оценивая остроту зрения по опознанию или по обнаружению изображений различной величины. Для оценки остроты зрения по опознанию применяют специально подобранные оптотипы — такие как кольца Ландольта, буквы. Для маленьких детей подобные тесты не применимы, для них разработаны специальные оптотипы в виде геометрических фигур или простых рисунков обычных предметов [таблица Орловой, тесты Lindstet, 1988; Hyvarinen, et al., 1992], но измерять остроту зрения по этим тестам возможно не ранее 1.5–2 лет. Наиболее простой способ оценки остроты зрения, доступный для малышей начиная с 4–5 месячного возраста, — по обнаружению мелких объектов (например, тесты Шеридана [Sheridan, 1970]), — определяется минимальный размер тех объектов, который малыш может фиксировать или следить за их движением.

Применение таких разнородных тестов может приводить к оценке функционирования различных пространственных элементов. Можно с большой долей уверенности утверждать, что обнаружение сигналов малых размеров обеспечивают наиболее чувствительные

пространственные элементы зрительной системы, настроенные на оптимальные средние частоты. (Необходимо заметить, что диапазон оптимальных частот зависит от возраста наблюдателя.) Опознание формы изображений типа колец Ландольта на пределе разрешения зрительной системы связано с функционированием самых высокочастотных элементов — «детекторов» полос и краев. Различение прямоугольных и синусоидальных решеток также происходит при участии самых высокочастотных элементов, но этими элементами могут быть и «детекторы» решеток. Использование же других изображений для оценки остроты зрения требует их специального рассмотрения: как минимум, применения к ним пространственно-частотного анализа и сопоставления результатов этого анализа с экспериментальными данными. Поэтому изменение показателей остроты зрения с возрастом при использовании разных тестов хотя и дает возможность оценивать развитие зрительной системы, но, скорее всего, при этом происходит оценка развития ее различных структур.

Для оценки остроты зрения у младенцев мы использовали задачу на обнаружение сигнала — определяли минимальный размер крупинок (шарики из черного хлеба на белом фоне), которые мог заметить малыш. [Подобная процедура оценки остроты зрения с небольшими вариациями описана также в монографии Хювяринен,1996]. Диаметр крупинок варьировал от 4.0 до 0.5 мм: 4, 3, 2, 1 и 0.5 мм. Несколько крупинок (обычно от 3 до 6) предъявляли на светлом фоне на расстоянии 25–30 см от глаз малыша справа или слева от направления его взора. Расстояние между крупинками значительно превосходило их размер. Яркость фона была равна 90 кд/м^2, крупинок — 5 кд/м^2.

Ребенок сидел на коленях у родителей или воспитателей перед столом. Экспериментатор находился напротив него. Он пытался привлечь внимание ребенка, чтобы тот посмотрел на тот участок стола перед собой, где справа или слева находились крупинки. Местоположение крупинок (справа или слева) менялось случайным образом, как правило, вместе с заменой их размера. Экспериментатор отмечал реакцию ребенка на предъявление крупинок, начиная процедуру с крупинок большого размера, постепенно заменяя их на меньший размер. Реакция детей могла быть разнообразной: старшие дети пытались

взять крупинки рукой, их потрогать или похлопать по ним, младшие дети могли просто дольше смотреть в ту сторону, где были крупинки, не трогая их руками. Если никакой реакции на крупинки данного размера не было замечено, пробу повторяли 2–3 раза, иногда после других тестов. Метод был скриниговый: подобная процедура занимала всего около 5–7 минут. По своей основе он близок методу предпочтительного смотрения. (Следует заметить, что при полном отсутствии реакции ребенка на данный тест смотрели его реакцию на изображения лица разного размера и другие более яркие или большие по размеру стимулы, чтобы грубо оценить уровень зрения ребенка).

Кроме того, мы обследовали и другие базовые зрительные функции. Мы приводим их краткое описание для того, чтобы показать обоснованность дальнейшего нашего вывода о связи именно данной оценки остроты зрения с выполнением других более сложных зрительных задач у детей младшего возраста. При обследовании других зрительных функций малышей мы также использовали скрининговые методы. При оценке бинокулярной фиксации и ее устойчивости отмечали, насколько симметрично отражается пламя свечи, на которую смотрит младенец, в зрачках обоих глаз. У детей старше 6 месяцев проверяли, сохраняется ли устойчивая бинокулярная фиксация при отсутствии специального фиксируемого стимула. Определяли границы поля зрения по горизонтали и симметричность его справа и слева: регистрировали момент фиксационного поворота глаз и головы младенца, когда игрушка, находящаяся первоначально за пределами его поля зрения, входила в поле зрения. Проверяли, выполняет ли малыш все типы прослеживающих движений глаз (горизонтальные, вертикальные, круговые и конвергенционно-дивергенционные), являются ли они плавными или скачкообразными, одинаково ли прослеживание при движении стимула в противоположных направлениях. Смотрели, одинакова ли реакция (сопротивление) малыша на прикрывание его левого и правого глаза (кавер-тест). Подробно описание процедуры тестирования и возрастные нормативы опубликованы в работе [Невская и др., 1996], близкие процедуры использовала и Хювяйринен [Хювяйринен, 1996].

У всех детей также оценивали общий уровень их психомоторного развития: на основе опросника, заполняемого родителями.

В опроснике (КИД-шкала, разработанная Рейтер [Reuter, 1985]) нужно было ответить, что умеет делать ребенок. Эта шкала содержит более 250 вопросов. Она достаточно широко применяется в скандинавских странах и в США. В нашей стране шкала также апробирована уже на достаточно большом количестве детей и по ней получены прикидочные нормативные данные. Для каждого ребенка рассчитывали коэффициент K — отношение возраста психомоторного развития малыша, определяемого по КИД-шкале, к его физическому возрасту. В случае полного соответствия возраста развития и физического возраста коэффициент K близок к 1, при опережающем развитии K>1, при отставании K <1. Кроме того, по этой же шкале, используя часть вопросов, определяли отдельно уровень когнитивного развития ребенка. Этот показатель мы также учитывали в своей работе для анализа данных.

Всего было обследовано более 1000 детей в возрасте от 4 до 18 месяцев, родители которых обратились в Центр раннего вмешательства для консультации по поводу развития их ребенка. Для определения возрастных нормативов развития остроты зрения использовались только данные, полученные на практически здоровых детях, не имеющих существенного отставания в психомоторном развитии по шкале KID [Reuter, 1985]. Контроль за состоянием здоровья детей осуществлялся с помощью врачей педиатров, неврологов и окулистов. Результаты исследований, описанные в настоящем параграфе, опубликованы в работах [Невская и др., 1996, 1998; Бондарко и др., 1999].

В табл. 2 приведены данные по измерению остроты зрения у детей в возрасте от 3 до 15 месяцев тестами на обнаружение мелких объектов. Во второй строке таблицы приведено распределения детей по возрастным группам. Далее указан процент детей, реагирующих на крупинки разного размера, который отражает возможность оценки остроты зрения таким методом у младенцев разного возраста. В следующей строке приведены средние значения минимально видимых крупинок, которые могли заметить малыши в каждом возрасте. Затем в строках таблицы приведено стандартное отклонение и ошибка измерения, вычисленная как отклонение, деленное на корень квадратный из количества измерений.

Таблица 2.

Измерение остроты зрения у детей в возрасте от 3 до 15 месяцев тестами на обнаружение мелких объектов.

Возраст детей (месяцы)	3–3.9	4	5	6	7	8	9	10	11	12	13–15
Число обследованных детей	25	61	51	109	79	87	74	78	41	37	71
% удачных измерений	32	55.7	75.5	96.3	94.9	98.3	98.7	100	100	98.6	100
Размер крупинок (мм)	1.5	1.56	1.36	0.78	0.58	0.52	0.53	0.51	0.51	0.52	0.53
Дисперсия	0.82	0.88	0.94	0.72	0.42	0.24	0.34	0.19	0.24	0.20	0.12
Ошибка	0.43	0.15	0.15	0.07	0.12	0.03	0.04	0.02	0.04	0.03	0.01

Данные детей старше 15 месяцев не показаны по причине неадекватности данных тестов для оценки остроты зрения у старших детей. Дело в том, что применение крупинок таких размеров фиксированной яркости ограничило наши измерения. Более мелкие крупинки практически невозможно было изготовить, поэтому точная оценка развития остроты зрения этим тестом была нами получена только для детей до 8-месячного возраста. Это видно из таблицы. Начиная с этого возраста, оценки остроты уже не увеличиваются. Размер крупинок при этом составляет около 0.5 мм, что при данном расстоянии до глаз ребенка соответствует размеру примерно в 5 угл. мин. или пространственной частоте 6 цикл/град.

Как видно из таблицы, средний минимально видимый размер крупинок в возрасте 3–4 месяцев составляет 1.5 мм (частота около 2 цикл/град.), затем уменьшается примерно в 3 раза к 8-месячному возрасту (частота около 6 цикл/град.). При этом только треть детей в возрасте от 3 до 4 месяцев обращает внимание на мелкие детали, в то время как практически все дети реагируют на крупинки в возрасте 6 месяцев.

Зависимость, полученная нами, достаточно хорошо соответствует данным других авторов, например, Ридберг [Rydberg, 1998], где острота зрения по обнаружению определялась по методике Шеридана [Sheridan, 1970]. В отличие от наших исследований в этой методике использовались катящиеся шарики разного размера. По срав-

нению с приведенными отдельными данными Хювяйринен (1996), использующей стационарные крупинки (сахар или драже разного цвета), у нас получены несколько меньшие размеры минимально видимых крупинок в возрасте 6 месяцев. Объясняться это может как расхождениями в условиях проведения экспериментов (не стандартизированная средняя яркость фона и крупинок, их цветовая гамма и т.п.), так и несколько различной методикой проведения экспериментаторов. У Хювяйринен экспериментатор, судя по описанию, пассивно наблюдал за поведением ребенка. Мы же старались привлечь его внимание: посмотреть на стол с крупинками.

Как было отмечено ранее, трудно ожидать, что подобным методом мы измеряли остроту зрения, которая связана с опознанием формы мелких объектов или с разрешающей способностью зрительной системы. Поэтому интересно сопоставить наши данные с результатами других исследователей, проведенных, например, с использованием синусоидальных и прямоугольных решеток. (Отметим, что для таких маленьких детей невозможно использовать стандартные тесты на опознание объектов).

При оценке остроты зрения с использованием решеток наблюдаются расхождения в оценках, зависящие от применяемого метода исследования. У новорожденных острота зрения по предпочтительному смотрению составляет около 0.6–1.0 цикл/град. и возрастает до 12–15 цикл/град. к концу первого года жизни. При этом особенно бурный рост остроты, измеренной этим способом, наблюдается в возрасте до 6–8 месяцев. В 6–8 месяцев острота оценивается примерно предельной частотой в 9–10 цикл/град., что не значительно отличаются от полученных нами показателей по определению предельной частоты с помощью крупинок различного размера. Регистрация вызванных потенциалов дает более высокие значения остроты зрения для новорожденных — от 2 до 4.5 цикл/град. и около 19 цикл/град. в возрасте 8 месяцев. Различия между значениями остроты зрения, измеренными этими методами, постепенно уменьшаются с возрастом. Нормативные данные об остроте, измеренной по предпочтительному смотрению или по вызванным потенциалам с использованием решеток у младенцев и детей разного возраста можно найти в обзоре [Teller, 1997]. Предположения о том, чем мо-

гут быть вызваны расхождения между данными, полученными при использовании различных методов, высказаны в нашей монографии [Бондарко и др., 1999]. Метод, примененный нами, аналогичен методу предпочтительного смотрения, поэтому логичнее сравнивать наши данные с результатами, полученными этим способом. При таком сопоставлении для восьмимесячных младенцев острота зрения отличается в полтора раза.

Улучшение остроты зрения в возрасте до года многие исследователи связывают главным образом с развитием сетчатки. Как показали исследования Фиорентини с соавторами [Fiorentoni et al., 1983], увеличение остроты зрения, определяемой по корковым вызванным ответам, идет параллельно развитию разрешающей способности сетчатки, определяемой по электроретинограмме. Следовательно, лимитирующий фактор, который определяет низкую остроту зрения в младенческом возрасте — это развитие сетчатки. Различия в остроте зрения у младенцев и у взрослых объясняются, в основном, хотя и не полностью, различиями в оптике глаза, в плотности фоторецепторов в фовеальной области и в их чувствительности [Banks, Bennet, 1988; Wilson, 1988; Teller, 1997]. Но нельзя и не учитывать при рассмотрении вопроса о развитии остроты зрения с возрастом состояние и центральных отделов зрительной системы. Наиболее полно это развитие отражено в изменениях, происходящих с возрастом в функциях контрастной чувствительности, измеренных с помощью синусоидальных решеток различных пространственных частот.

Контрастная чувствительность, определяемая как по предпочтительному смотрению, так и по вызванным потенциалам, оказывается у младенцев значительно ниже, чем у взрослых, причем различия особенно велики для высоких пространственных частот. Вся функция контрастной чувствительности у младенцев сдвинута в область низких частот, причем в этой функции у младенцев моложе 2.5 месяцев в отличие от взрослых нет спада чувствительности для частот ниже 1 цикла на градус. С возрастом контрастная чувствительность, как и острота зрения, улучшается. Наиболее резко контрастная чувствительность возрастает в первые месяцы жизни, а затем — более медленно. Чувствительность к низким пространственным частотам достигает взрослого уровня уже к концу первого года жизни ребен-

ка, а чувствительность к высоким частотам — только к 8–10 годам [Atkinson et al., 1974, 1977; Banks, Salapatek, 1981, 1983; Norcia et al., 1990].

Функцию контрастной чувствительности принято рассматривать как общую огибающую для чувствительности нескольких относительно независимых друг от друга пространственно-частотных каналов, существующих в зрительной системе. Как говорилась выше, каждому каналу соответствует набор пространственных элементов, настроенных на частоту канала. Этими пространственными элементами могут быть описанные нами ранее «детекторы» полос, краев и решеток. Изменения в пространственно-частотных каналах с возрастом неразрывно связаны с изменениями в элементах, их образующих. В сформировавшейся зрительной системе при различении ориентации высокочастотных решеток и обнаружении мелких объектов функционируют различные пространственные элементы (см. разделы 1–3). С этой точки зрения близкие оценки остроты, полученные в этих задачах у младенцев, могут объясняться тем, что в первый год жизни описание пространственных свойств объектов обеспечивает малое количество пространственных элементов [Wilson, 1988], настроенных только на низкие частоты. Возможно, даже, что этими двумя методами мы оцениваем развитие одного и того же частотного канала, поскольку каналы в этом возрасте широкополосные [Wilson, 1988]. Несмотря на то, что контрастная чувствительность к низким частотам практически не увеличивается с возрастом, начиная с годовалого возраста, мы не можем утверждать, что после этого возраста не происходит изменений в организации пространственных элементах, образующих соответствующие пространственно-частотные каналы. Как было сказано в разделе 1, для зрительной системы взрослого наблюдателя имеет место согласованность в размерах и частотной настройке среднечастотных и высокочастотных пространственных элементов, которая необходима для успешной совместной обработки различных зрительных сигналов.

Проследить согласованность формирования элементов, настроенных на разные частоты, в онтогенезе, насколько нам известно, пока не удавалось. В литературе имеется только небольшое количество описаний пролонгированных исследований, выполненных на

одних и тех же детях в разном возрасте. К ним относится сборник статей под редакцией Ридберг [Rydberg, 1998], в котором опубликованы материалы ее сообщений с соавторами, доложенные на международной конференции членов Ассоциации исследований в области зрения и офтальмологии (ARVO — 1994) и по «сниженному» зрению –“International conference on low vision”, Мадрид, 1996. Мы приведем тут достаточно подробное описание этих работ, поскольку они подтверждают наше предположение об использовании зрительной системой разных пространственных элементов, к тому же, видимо, по-разному формирующихся в онтогенезе.

В разных статьях сборника [Rydberg, 1998] представлены данные детей с нормальным зрением и различными зрительными заболеваниями (в отдельную группу выделены дети с косоглазием) в возрасте от 1 месяца до 1.5 лет и от 1.5 до 6 лет, а также взрослых людей с нормальным зрением и с косоглазием. У детей младшего возраста (до 1.5 лет) острота зрения определялась двумя методами: по различению решеток с помощью карт Теллер (прямоугольные решетки различных пространственных частот) методом предпочтительного смотрения и по обнаружению сигналов по методике Sheridan [Sheridan, 1970] с помощью движущихся объектов малого размера (Stycar-rolling balls). У детей старшего возраста, начиная с полутора лет, и у взрослых острота зрения определялась, кроме того, и тестами на опознание (авторы использовали предметные изображения или буквы латинского алфавита). В результате исследований было показано следующее. У групп детей в возрасте до 1.5 лет с нормальным зрением и с различными глазными заболеваниями не было получено достоверных отличий по остроте зрения, определенной по различению и обнаружению сигналов с помощью решеток и крупинок. И в той, и в другой группе наблюдался большой разброс этого показателя. На основании этих данных авторы делают вывод о том, что тесты по обнаружению и различению не могут быть использованы для диагностических целей. Но они могут быть полезны для общей оценки зрительной функции. Только тесты на опознание, возможность применения которых ограничена возрастом детей, дают адекватную картину — показывают низкую остроту при зрительных заболеваниях. Далее, в исследовании [Rydberg, 1998] не было

выявлено корреляции между сниженной остротой в малом возрасте, определенной по картам Теллер, и в большом возрасте (4–6 лет), определенной тестами на опознание изображений.

С нашей точки зрения, отсутствие такой корреляции не кажется странным по нескольким причинам. Во-первых, как было сказано ранее, у детей малого возраста функционируют только низкочастотные каналы, высокочастотные же еще не сформированы. Во-вторых, зрительная система может использовать разные РП или пространственные элементы для различения решеток и опознания предметных изображений. Поэтому может не быть корреляции между оценками остроты даже при использовании одного и того же тестового стимула. В приведенных Ридберг данных, настораживает тот факт, что при использовании карт Теллер все взрослые наблюдатели и старшие дети с нормальным зрением — в работе приводятся данные более 40 человек — имеют остроту 60 цикл/град. — самую высокую, которую возможно получить при использовании данных тестов. То есть, возникает естественный вопрос о корректности данных тестов, поскольку частота 60 цикл/град. слишком велика для среднего наблюдателя с нормальным зрением. Сами авторы обращают на этот факт внимание и говорят, что это может быть связано с тем, что наблюдатель выполняет задачу обнаружения края решетки, а не различения решетки от гомогенного поля с той же яркостью. (Напомним, что карты Теллер представляют собой прямоугольные решетки, которые отделяются от среднего фона полосой, ширина которой в 2 раза меньше ширины полосы решетки). На такую возможность замены зрительной задачи при использовании тестов Теллер впервые указали Робинсон и др. [Robinson et al., 1988, цит. по: Rydberg, 1998]. Однако Ридберг и соавторы [Rydberg, 1998] пишут, что, скорее всего, не происходит замены задачи. Они обосновывают свой вывод тем, что Добсон и Луна [Dobson, Luna, 1993, цит. по: Rydberg, 1998] получили одинаковую остроту зрения у детей 4–36 месяцев при использовании обычных карт Теллер и решеток с измененной круглой апертурой. При этом во втором случае сравнение шло не со средним фоном, а с высокочастотной решеткой, окруженной специально отфильтрованной апертурой. Далее, для подтверждения своего предположения об игнорировании края

у карт Теллер авторы в качестве аргумента приводят следующий факт. Они пишут, что в более ранних исследованиях Ридберг при использовании круглой апертуры у решеток и процедуры предпочтительного смотрения у амблиопичного глаза получается переоценка остроты при использовании решеток по сравнению со стимулами на опознание, в то время, как такой эффект отсутствует у неамблиопичного глаза тех же наблюдателей. Это, по мнению автора, свидетельствует в пользу того, что дело тут не в апертуре, а в самих стимулах.

На наш взгляд, приведенные Ридберг факты не противоречат тому, что исходные карты Теллер могут быть некорректными для взрослых наблюдателей с нормальным зрением, поскольку зрение у детей и взрослых существенно различается. Возможное объяснение расхождений данных у детей и взрослых при применении карт Теллер может заключаться в том, что у детей младшего возраста задача обнаружение края может совпадать с задачей различения решеток из-за отсутствия разнообразия частотных каналов. К сожалению, Ридбер не проводит сопоставления результатов тестирования по обнаружению и различению сигналов у детей младшего возраста, имея в наличие эти данные. Правда, на приведенных в работах Ридберг рисунках в возрасте до 8 месяцев фигурируют данные только нескольких детей, поэтому такое сопоставление, возможно, не было бы эффективным. В более же позднем возрасте у нее не получено увеличения остроты зрения на обнаружение мелких объектов, как и у нас, из-за чисто технических причин — не было объектов еще меньшего размера.

Расхождения в оценках остроты между амблиопичным и неамблиопичным глазами кажется заслуживающим внимания с точки зрения сопоставления примененных тестов. Возможное объяснение различия оценок кроется как раз в том, что для амблиопичного глаза не могут быть сформированы высокочастотные элементы, описывающие предметные изображения, называемые нами «детекторами» полос и краев. Дело в том, что эти высокочастотные элементы (в нашем исследовании это показано для «детекторов» полос) устроены таким образом, что у них наблюдается антагонизм тормозных и возбудительных on и off зон, а размеры зон совпадают

с шириной функции рассеяния и с размером гексагона. Задачу же различения решеток могут выполнять и другие элементы — «детекторы» решеток, для которых подобное согласование не обязательно, поскольку они могут быть образованы только возбудительными субполями.

Анализ данных, проведенный Ридберг, показывает, что не прослеживается зависимости между сниженной остротой, определенной у детей до 1.5 лет и в старшем возрасте используемыми для детей этих возрастных групп тестами, наиболее распространенными при подобных обследованиях. Возможно, это связано с неадекватным подбором тестов. Для детей самого младшего возраста для определения их зрительной функции используют изображения лица разного размера и контраста или же отфильтрованные изображения лица, так как дети реагируют на эти стимулы, начиная с самого рождения. Несмотря на то, что определение остроты зрения в самом раннем возрасте описанными выше методами не может служить для диагностики зрительных заболеваний и являться прогнозом для выявления каких-либо отклонений в зрительной функции в дальнейшем, эту процедуру используют для оценки зрительной функции детей самого малого возраста. В наших исследованиях мы выявили корреляцию между сниженной остротой зрения и выполнением различных зрительных задач [Невская и др., 1998]. Проведенные исследования подтвердили предположение, что при подобном методе определения остроты зрения мы оцениваем функционирование наиболее чувствительных каналов зрительной системы. К сожалению, мы практически были лишены возможности провести пролонгированные исследования с детьми со сниженной остротой зрения. Поэтому не можем сказать, как сниженная острота сказалась в дальнейшем на общем развитии и развитии зрительной функции у детей.

4.2. Измерение остроты зрения и краудинг-эффекта у детей и подростков в возрасте от трех до семнадцати лет

В этом разделе мы остановимся на исследованиях, которые важны для описания формирования нейрофизиологических механизмов краудинг-эффекта в онтогенезе, а также для офтальмологической практики.

Известно, что острота зрения у детей в возрасте примерно до 6 лет ниже, чем у взрослых [см. например, Хювяйринен, 1996]. Однако обычно при измерении остроты зрения у дошкольников используются таблицы, где расстояния между символами в строке равны ширине символов. Поэтому многие исследователи считают, что низкая острота у детей младшего возраста является следствием влияния на опознание тестового стимула окружающих объектов, то есть, более выраженного краудинг-эффекта.

В обзоре Аткинсон [Atkinson, 1991] описаны различия, полученные на детях и взрослых при изучении краудинг-эффекта. Показано, что различное окружение по-разному воздействует на восприятие центрального стимула у детей и взрослых. Но в большинстве случаев на одинаковом расстоянии окружения по отношению к центральному стимулу дети 5-летнего возраста дают меньший процент правильных ответов, чем 7-летние дети или взрослые. При сравнении исследований краудинг-эффекта, проведенных на взрослых и детях, использовались стимулы с фиксированным расстоянием до окружения. Как правило, были взяты расстояния, кратные 1/2 размера центрального стимула (такие стимулы подобны фрагменту таблиц, используемых для определения остроты зрения). В этих условиях на взрослых было получено, что при расстоянии, равном размеру стимула, ухудшения в восприятии по отношению к изолированному стимулу не наблюдается. В то же время, при таком же расстоянии дети 5-летнего возраста видят центральный стимул хуже, чем аналогичный изолированный стимул. Примерно в 7 лет дети начинают воспринимать такие стимулы аналогично взрослым. Расстояния большие, чем размер стимулов, в известных нам исследованиях не использовались. Поэтому по результатам этих исследований невозможно судить о размере тормозного взаимодействия в краудинг-эффекте у детей различного возраста.

Гипотезы по поводу худшего опознания центральных стимулов у детей по сравнению со взрослыми высказывались следующие [см. Atkinson, 1991]: 1) по сравнению со взрослыми детям изначально труднее объяснить исходную задачу и сложнее получить от них словесный отчет, 2) у детей недостаточно внимания для решения такой задачи. При этом Аткинсон считает, что сам механизм краудинг-эф-

фекта и у детей, и у взрослых заключается, главным образом, в тормозном воздействии друг на друга близлежащих контуров при их зрительном восприятии [Woodrow, 1938]. Автор не высказывает предположение о том, каким механизмом обеспечивается это тормозное воздействие.

В настоящем исследовании мы хотим не только показать, что у детей происходит ухудшение восприятия тех стимулов, на которых у взрослых оно не наблюдается, но и посмотреть изменение размера тормозного взаимодействия в зависимости от возраста. Тем самым, рассмотреть механизм формирования краудинг-эффекта в онтогенезе. С этой целью мы использовали различные расстояния до окружения, соответствующие 0.75–2.0 размерам тестового стимула. В качестве тестовых стимулов были взяты кольца Ландольта, а окружения — полосы. Исследования восприятия окруженных стимулов на пределе разрешения зрительной системы были проведены также для сравнения в тех же условиях на группе взрослых наблюдателей. В этом случае расстояния до окружения менялись в пределах 0.25–2.0 размеров тестового стимула.

Сама процедура изучения краудинг-эффекта, как отмечалось ранее, протекает следующим образом: вначале определяется минимальный размер для тестовых стимулов, при котором эти стимулы еще видны (процент правильных ответов выше случайного уровня), а затем изучается влияние окружения на опознание стимулов этого минимального размера. Таким образом, при определенном выборе тестовых стимулов одним из этапов проведения такого исследования может являться определение остроты зрения.

Острота зрения является важным показателем развития всей зрительной системы в целом и оценки ее функционального состояния. (Это продемонстрировано и в предыдущем параграфе.) Для того чтобы рассмотреть отдельно показатели остроты зрения и краудинг-эффекта, в настоящем исследовании мы в одном эксперименте измеряли остроту зрения на изолированных стимулах и изучали восприятие аналогичных стимулов при окружении их дополнительными контурами.

Для тестирования были разработаны программы для ПЭВМ. В качестве стимулов использовались кольца Ландольта, которые

предъявлялись в центре экрана с разрывом в одной из 4-х ориентаций (вверху, внизу, справа или слева). Стимулы синтезировались на стандартном цветном дисплее с размером экрана 14 дюймов по диагонали. Разрешение экрана 480 на 640 пикселов с размером пиксела 0.29 мм, яркость фона была равна 80 кд/м^2, яркость стимулов — 5 кд/м^2.

Процедура определения порогового размера кольца заключалась в следующем. Вначале экспериментатор, показывая кольца разного размера, приблизительно подбирал минимальный размер кольца, при котором ребенок еще хорошо видел его разрыв. Для точного определения порогового размера кольца использовался метод лестницы. В центре экрана предъявляли кольцо Ландольта, начиная с размера, определенного экспериментатором. Если при предъявлении четырех ориентаций кольца данного размера ребенок давал не менее трех правильных ответов, то в следующей серии размер стимула уменьшался на 0.1 лог.ед. В противном случае размер увеличивался на такую же величину. Для определения порогового размера усредняли минимальные и максимальные значения размеров, полученные в точках перехода от правильных к неправильным ответам. Такой способ соответствует определению 75 % уровня правильных ответов.

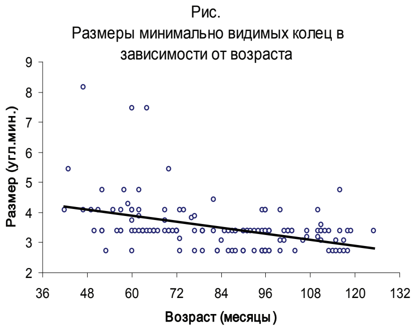
В дальнейшем при изучении краудинг-эффекта ребенку предъявлялся тестовый стимул определенного таким способом размера, окруженный полосами с 4-х сторон (рис. 5). Расстояние до полос варьировали случайным образом в пределах от 0.75 до 2 размеров кольца с шагом, равным 0.25 размера кольца. Вся процедура определения минимального размера и исследования краудинг-эффекта при 6-ти расстояниях до окружения и предъявлении колец Ландольта по одному разу в каждой из 4-х ориентаций занимала около 15 минут (время примерно одинаково распределялось на исследование краудинг-эффекта и определение минимального размера колец). Как правило, с каждым ребенком проводилось 2–4 повторных исследований в тот же день и еще через несколько дней. Для подсчета усредняли результаты нескольких экспериментов, при этом не учитывались нестабильные первые пробы.

Экспериментатор просил ребенка определить ориентацию кольца Ландольта. Ребенок или показывал рукой, где находится разрыв, или ему в руки давали объемное эталонное изображение кольца (типа бублика или руля с вырезанной одной стороной), он поворачивал этот эталон, совмещая разрыв с видимым разрывом в кольце. Большие дети давали словесный ответ. Для того чтобы ребенок в ходе эксперимента не уставал, исследование проводилось в игровой форме. Наблюдение было бинокулярным с расстояния 4.3 м до экрана. При таком расстоянии размер одного пиксела на экране равен 0.34 угл. мин.

У всех детей предварительно была оценена острота зрения по таблице Орловой. Все взятые в рассмотрение дети имели нормальную (не ниже 1) остроту зрения и не имели каких-либо глазных заболеваний. Взрослые наблюдатели также имели остроту зрения не ниже 1, измеренную по таблице Головина-Сивцева. Ранее они не принимали участие в каких-либо психофизических экспериментах. В таблице 3 показано число обследованных детей и взрослых в зависимости от их возраста.

Были получены размеры колец, соответствующие 75-процентному уровню правильных ответов при определении ориентации разрывов в кольцах Ландольта при использовании стимулов без окружения. Средние величины этих размеров в зависимости от возраста приведены в табл. 3. На рис. 10, А показаны результаты измерения предельно видимых размеров колец отдельно для каждого ребенка в зависимости от его возраста. Возраст детей дан в месяцах. По этим данным построена линия регрессии. Корреляция равна 0.46, достоверность $p < 0.001$. Таким образом, можно утверждать, что острота зрения увеличивается с возрастом у детей выбранной возрастной группы. Для определения возраста детей, при котором острота зрения перестает улучшаться, мы разбили детей по возрастным группам в пределах одного года и посчитали в каждой группе среднее и стандартное отклонение. Результаты этих расчетов показаны на рис. 10, Б и приведены в таблице 3. Из рисунка видно, что острота зрения практически не улучшается и стабилизируется где-то в возрасте 7–8 лет.

А



Б

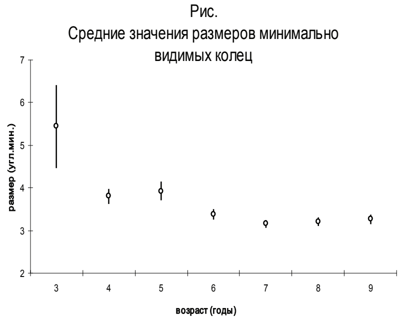


Рис. 10. Размеры минимально видимых колец Ландольта в зависимости от возраста детей.

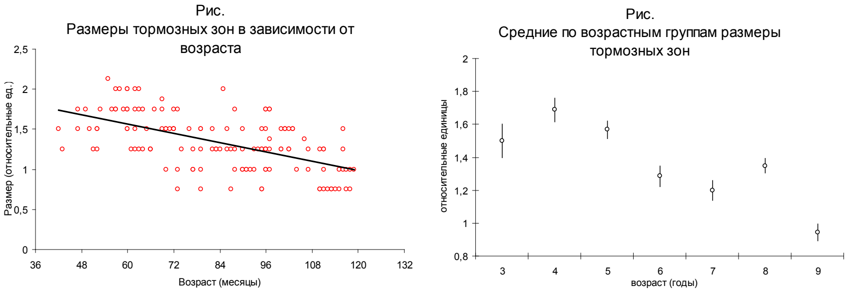
А — индивидуальные данные, *Б* — средние значения в каждой возрастной группе

Для анализа данных по исследованию краудинг-эффекта для каждого наблюдателя были построены кривые процента правильных ответов в зависимости от расстояния до окружения. Практически все кривые монотонно возрастали, их значения при увеличении расстояния приближались к значениям правильных ответов на тестовый стимул без окружения

Для каждого наблюдателя был определен размер тормозной зоны. Из-за сравнительно небольшого числа накоплений (более 20 для каждого расстояния до окружения) мы не могли применить строгие критерии для нахождения значений зон (например, по методу Хи-квадрат как достоверную разницу по сравнению с ответом на изолированный стимул). Проводилась визуальная экспертная оценка кривых. Такая оценка возможна, так как кривые имели характерный вид: выходили на плато при достижении определенного расстояния между окружением и тестовым стимулом. Расстояния, предшествующие выходу кривых на плато были взяты за размеры тормозных зон.

На рис. 11, А приведены размеры тормозных зон для каждого наблюдателя в зависимости от его возраста. По оси абсцисс отложен возраст каждого ребенка в месяцах, по оси ординат — размер тормозной зоны, выраженный в относительных единицах как рас-

стояние до окружения в угловых минутах перед выходом кривой на плато, деленное на размер минимально видимого кольца. На этом же рис. показана линия регрессии. Корреляция равна 0.62, что соответствует 0.001 уровню достоверности. Видно уменьшение тормозных зон с увеличением возраста детей. В табл. 3 приведены средние значения размеров тормозных зон, выраженные в относительных единицах как абсолютные размеры, деленные на размер кольца Ландольта, в каждой возрастной группе, и их стандартные отклонения. Эти величины показаны на рис. 11, Б. На этом рис. показано постепенное уменьшение размера тормозной зоны в зависимости от возраста.



А Б

Рис 11. Расстояния между тестовыми кольцами Ландольта и полосами, при которых полосы влияли на опознание ориентации колец.

А — индивидуальные данные, *Б* — усредненные по возрастным группам

В табл. 3 также приведены данные по оценке остроты зрения и размера тормозных зон у взрослых людей, полученных в аналогичных условиях. Основную группу наблюдателей составили студенты университета СПБГУ в возрасте 19–22 лет. Острота зрения, измеренная по таблице Головина-Сивцева, у всех наблюдателей была равна 1 или выше. Размер тормозной зоны у всех взрослых наблюдателей был приблизительно равен размеру кольца. Кроме того, в таблице приведена корреляция между оценками остроты зрения и размерами тормозных зон. Ее значение равно 0.22. Это значение является достоверным на уровне значимости $p = 0.05$.

Таблица 3

Зависимость размеров минимально видимых колец Ландольта и тормозных зон от возраста наблюдателей

Возраст наблюдателей	Размер колец (угл.мин)	Стандартное отклонение	Размер зоны (отн.ед.)	Стандартное отклонение	Количество наблюдателей
3–4 г.	5.4	1.9	1.4	0.15	4
4–5 л.	4.4	1.8	1.7	0.11	5
5–6 л.	3.9	1.2	1.7	0.23	11
6–7 л.	3.7	0.4	1.3	0.31	9
7–8 л.	3.1	0.6	1.2	0.25	27
8–9 л.	3.1	0.7	1.3	0.30	19
9–10 л.	3.3	0.5	0.9	0.26	21
Взрослые, средний возраст — 21 год	2.8	0.2	0.92	0.11	19
Корреляция между размером колец и размером зон равна 0.22					

Как видно из таблицы, острота зрения у детей сначала возрастает и становится близкой к взрослому уровню в возрасте около 7 лет, а размер тормозной зоны становится подобным взрослому уровню только примерно к 9 годам.

В результате обследования детей в возрасте от 3 до 9 лет получены данные по оценке остроты зрения и восприятию стимулов, окруженных дополнительными контурами. Показано, что острота зрения увеличивается с возрастом и примерно к 7-ми годам становится подобной взрослому уровню. Увеличение остроты зрения с возрастом наблюдали многие исследователи [Хювяринен, 1996], изучавшие эту функцию с помощью других тестов.

В офтальмологической практике за норму принято зрение, когда наблюдатель видит разрыв с шириной в 1 угл.мин. в кольце Ландольта размером в 5 угл.мин. У нас же, как и у других исследователей, средние групповые значения размеров колец для детей большого возраста были достоверно ниже 5 угл. мин. Аналогичные данные по оценке остроты зрения с помощью колец Ландольта получены в исследовании Г. И. Рожковой и др. [Рожкова, 2001], где показано, что острота зрения возрастает до значения, равного 1.9, в возрасте

9 лет. Такие высокие оценки остроты зрения, большие 1, полученные в экспериментах, могут быть связаны со спецификой использованных оптотипов, а не с исключительно хорошим зрением наших наблюдателей. В разделе 2 мы показали, что такое несоответствие оценок может быть устранено, если оценку остроты зрения рассматривать не как величину, обратную размеру разрыва у этих оптотипов, а как предельно видимую частоту в спектре этих изображений. Дело в том, что определять ориентацию кольца Ландольта тренированный наблюдатель может не по четко видимому разрыву, а по другому признаку — несимметричности изображения кольца: кольцо Ландольта при малых размерах имеет вид круглого пятна с уплощением с той стороны, где находится разрыв. В этом случае для опознания формы изображения достаточно более низких частот, меньших, чем частота разрыва примерно в 2 раза. Смена критерия испытуемым может являться основой для его обучения и наблюдаемого улучшения остроты зрения в ходе тренировки.

В наших экспериментах мы наблюдали увеличение остроты зрения и уменьшение тормозной зоны после первых предъявлений стимулов. В таблице приведены величины, когда эти параметры уже стабилизировались.

Важно отметить тот факт, что мы получили увеличение остроты зрения на изолированных стимулах, а не на стимулах с окружением, подобных стандартным таблицам из символов, применяемых для определения остроты зрения. Поэтому наши данные нельзя объяснить тем, что увеличение остроты зрения связано с более сильным влиянием окружения на опознания тестового стимула у детей младшего возраста. Кроме того, в наших исследованиях показано, что тормозная зона, в которой происходит ухудшение восприятия центрального стимула, также уменьшается с возрастом. Причем уменьшение зон не только абсолютное, которое может быть связано с увеличением остроты зрения, а относительное, то есть деленное на размер минимально видимого объекта. Размер зоны становится подобен взрослому уровню примерно только к 9 годам. Невысокая корреляция (см. таблицу), с нашей точки зрения, между оценками остроты зрения и размерами тормозных зон может служить основанием для рекомендации практикующим офтальмологам в использовании как

изолированных, так и окруженных стимулов при оценке функционального состояния зрительного аппарата у детей младшего возраста. Такая корреляция свидетельствует о слабой зависимости между механизмами, связанными с формированием остроты зрения и выделением объектов из окружения. Видимо, на оценку остроту зрения большее влияние оказывают возбудительные процессы, происходящие в нервной системе, тогда как на формирование тормозных зон, естественным образом, больше сказываются тормозные процессы.

В разделе, посвященном изучению краудинг-эффекта у взрослых людях, нами было показано, что краудинг-эффект с использованием тестовых колец Ландольта может быть объяснен функционированием самых высокочастотных пространственных элементов, образующих высокочастотный канал. (При использовании окружающих симметричных полос можно непосредственно описать ВФ «детектора» полосы.) Нейрофизиологическими коррелятами этих каналов могут являться рецептивные поля стриарной коры. Из морфологии известно, что формирование стриарной коры заканчивается к семи- девяти- летнему возрасту [Huttenlocher et al., 1982, см. также сопоставление морфологических и психофизиологических данных, Фарбер и др., 1990]. Таким образом, наши данные подтверждают предположение о локализации механизма краудинг-эффекта именно в стриарной коре. Напомним, что ранее Фломом и др. [Flom et al., 1963b] при предъявлении стимула одному глазу, а окружения другому было показано, что механизм краудинг-эффекта локализован на уровне не ниже стриарной коры.

Мы показали, что вне зависимости от изменения размера минимально видимого тестового кольца с увеличением возраста сужается тормозная зона взаимодействия, то есть, уменьшается размер тормозной зоны самого высокочастотного элемента. На языке пространственно-частотных каналов такое изменение означает сужение частотной настройки самого высокочастотного канала. Вилсон [Wilson, 1988] рассмотрел модель развития частотных каналов в раннем онтогенезе. В ней предполагается сужение ширины пропускания каналов с увеличением возраста. В данном исследовании мы подтвердили справедливость этого предположения для функционирования зрительной системы в более позднем возрасте.

Таким образом, в результате наших исследований мы получили изменения самого высокочастотного элемента зрительной системы- «детектора» полосы, у которого сужается и возбудительная зона и тормозная, но тормозная зона сужается в большей степени, что свидетельствует о сужении настройки самого высокочастотного канала. На рис. 12 показано изменение высокочастотного элемента — детектора полосы — в зависимости от возраста наблюдателей. Размер возбудительной зоны элемента у детей в возрасте 3–5 лет примерно в полтора раза больше, чем в 7–10 лет. Размер же тормозной зоны больше почти в 3 раза.

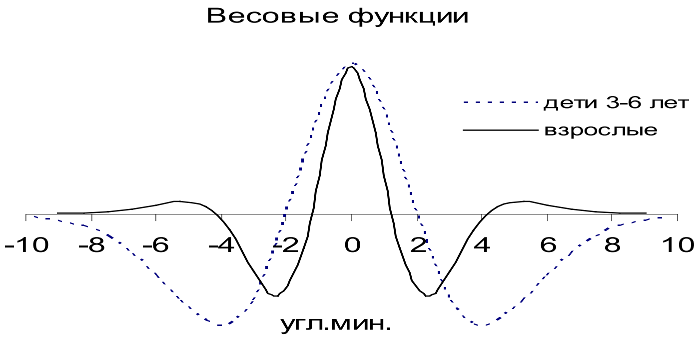


Рис. 12. Весовые функции самых высокочастотных элементов у взрослых и детей 3–6 лет.

Развитие пространственно-частотных каналов можно проанализировать и с другой точки зрения. Рис. 13 демонстрирует изменения в самых высокочастотных каналах, происходящие в онтогенезе. В основе построения графика лежат данные по остроте зрения, определенные методом предпочитаемого разглядывания с использованием прямоугольных решеток, приведенные в обзорной лекции Теллер [Teller, 1997], и наши данные по определению остроты с помощью колец Ландольта и прямоугольных решеток [Семенов и др., 2000; Бондарко, Семенов, 2005, 2012]. Здесь приняты в расчет значения остроты зрения у детей без патологии в зрительных функциях. На графике с увеличением возраста до семи-девяти лет предельная пространственная частота возрастает, а затем стабилизируется.

Исходя из такой зависимости, можно предположить, что младенец в момент рождения видит изображения размытыми, лишенными мелких деталей. С возрастом зрение существенно меняется, и только к семи годам становится возможным четкое видение изображений малых размеров таких, например, как мелкий шрифт текста. Заметим, что у наблюдателей без патологий старше семи лет средние показатели остроты зрения и предельные частоты выше, чем 1 ед. и 30 цикл/град.

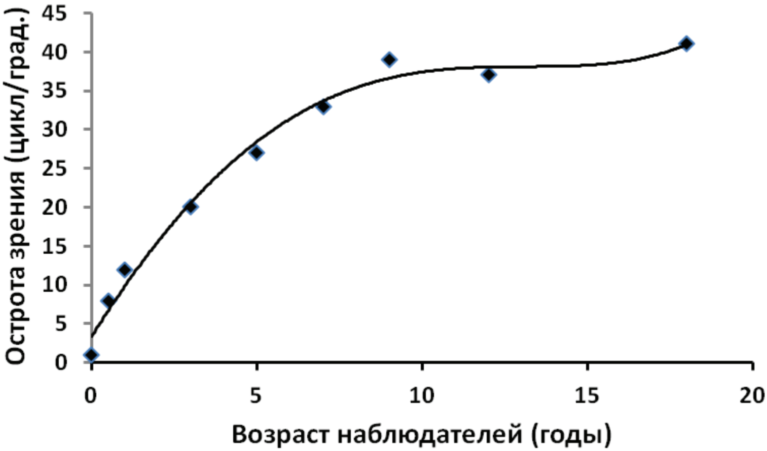


Рис. 13. Зависимость предельно видимых пространственных частот от возраста наблюдателей

4.3. Краудинг-эффект у детей при использовании различных оптотипов

Получив разные по размеру зоны у взрослых наблюдателей при использовании тестовых колец Ландольта и прямоугольных решеток (см. рис. 6) мы провели аналогичные исследования у школьников различного возраста. Заметим, что мы не могли провести аналогичные исследования у детей дошкольного возраста из-за сложности задачи. Тестовыми стимулами служили прямоугольные решетки в горизонтальной или вертикальной ориентациях, кольца Ландольта и буквы Е, повернутые вверх, вниз, вправо или вле-

во. При исследовании характеристик краудинг-эффекта решетки и буквы Е были окружены четырьмя аналогичными изображениями, расположенными на одинаковом расстоянии по горизонтали и вертикали относительно центрального тестового стимула. Ориентация каждого из четырех окружающих изображений менялась случайным образом от предъявления к предъявлению. Для колец Ландольта в качестве окружения были выбраны полосы фиксированной ориентации и той же ширины, что и кольца. Расстояние до центрального стимула менялось в ходе эксперимента. Примеры применявшихся в экспериментах стимулов показаны на рис. 5. Было обследовано 292 школьника 8–17 лет. Оценки остроты зрения, полученные тремя различными оптотипами показаны на рис. 4 в разделе 2.

На рис. 14 показаны средние величины максимальных расстояний до окружения, при которых ухудшается опознание тестовых стимулов, в зависимости от возраста наблюдателей (Бондарко, Семенов, 2005). По оси абсцисс отложен возраст в годах, по оси ординат — расстояние до окружения, выраженное в угл.мин. Виден различный характер зависимостей. Расстояния сначала уменьшаются, а затем стабилизируются для колец Ландольта и букв Е. Кривые для этих тестовых стимулов практически совпадают. Для решеток же происходит постепенное уменьшение критических расстояний с увеличением возраста. Линейная регрессия достоверно (p=0,05) приближает зависимость расстояний от возраста в случае использования тестовых решеток. Попарное сравнение средних величин по критерию Стьюдента выявило значимые различия (на уровне 5 %) при использовании букв Е у наблюдателей 8–9 и 12 лет. Также есть достоверные различия в средних у данных учащихся 8–9 лет и 16–17 при использовании тестовых решеток и у 12-летних наблюдателей при использовании тестовых решеток и букв Е. Таким образом, статистический анализ подтверждает наличие различных зависимостей от возраста при использовании тестовых решеток и букв Е или колец Ландольта. Необходимо отметить, что полученные размеры зон были меньше, чем в разделе 3, поскольку тут производилась более грубая оценка — метод был скрининговым.

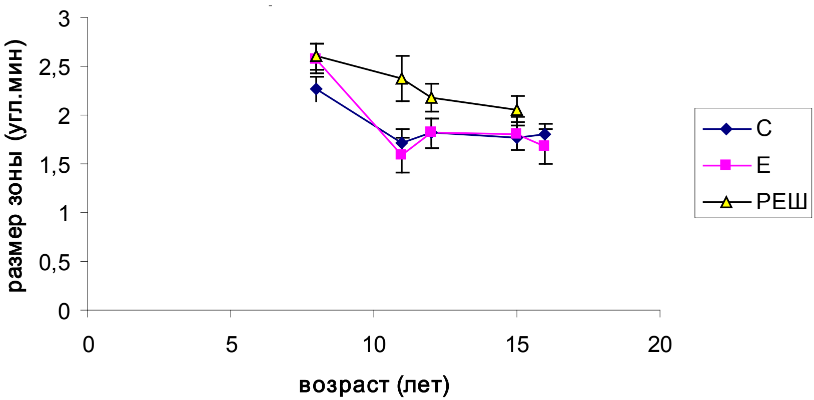


Рис. 14. Возрастная динамика изменения расстояния до окружения, при котором оно влияет на опознание тестовых стимулов, для разных тестовых стимулов и окружающих их изображений. Тестовые кольца Ландольта (С) окружены полосами, знаки Снеллена (Е) окружены аналогичными изображениями, прямоугольные решетки — решетками (РЕШ)

Таким образом, расстояния между стимулом и окружением, при которых происходило ухудшение, уменьшались с увеличением возраста для решеток до 16–17 лет, а для колец Ландольта и букв Е Снеллена до 12 лет. Могут быть высказаны различные гипотезы, объясняющие такие расхождения, учитывающие функционирование различных уровней зрительной системы. На низшем уровне такой характер зависимостей может свидетельствовать о более позднем формировании РП детекторов решетки, чем детекторов полос и краев. Возможно, что в данной задаче при использовании в качестве стимулов решеток проявляется механизм взаимодействия между пространственно-частотными каналами. Более вероятным представляется предположение, что при использовании решеток труднее происходит процесс выделения объекта из окружения, задействовано внимание, которое ослаблено у подростков.

Подведем итог представленным на рассмотрение исследованиям.

Измерение остроты зрения с помощью изолированных колец Ландольта у детей в возрасте от 3-х до 9-ти лет показало, что острота зрения увеличивается с возрастом и становится подобной взрослому уровню примерно к 7 годам. На измерение остроты зре-

ния у детей младшего возраста могут оказывать сильное влияние окружающие тестовые стимулы зрительные объекты. Расстояния, при которых окружающие изображения ухудшают восприятие тестовых объектов, уменьшаются с возрастом. Размер тормозной зоны в 3–6 лет превосходит размер минимально видимого кольца Ландольта в 1.7 раз и только к 9–10 годам становится таким же, как у взрослых наблюдателей. Проведенное исследование опознания тестовых стимулов, окруженных другими объектами, на детях разных возрастных групп позволяет сделать вывод о том, что формирование механизма краудинг-эффекта для оформленных стимулов (колец Ландольта и букв Е) заканчивается примерно к 9–12-ти летнему возрасту. Для текстурированных объектов этот процесс завершается позже.

В настоящем разделе представлены данные по изменению остроты зрения в зависимости от возраста детей при использовании разных тестов. Подтверждено наше предположение о том, что в этом случае остроту зрения обеспечивают различные пространственные элементы. Показана важность измерения этого показателя любыми тестами у детей до года, поскольку нормальное развитие детей сопровождается нормальной остротой зрения: дети со сниженной остротой отстают в когнитивном развитии от детей с нормальной остротой зрения, измеренной на обнаружение мелких объектов.

Для детей и подростком в возрасте от 3 до 17 лет представлены данные по измерению краудинг-эффекта. В результате показано, каким образом происходит формирование самых высокочастотных элементов с увеличением возраста детей. С нашей точки зрения, наблюдается слабая зависимость между изменениями возбудительных и тормозных зон элементов в онтогенезе. Это свидетельствует о необходимости сопоставления измерений остроты зрения у детей с помощью изолированных и окруженных другими изображениями стимулов. В разделе 1 показана, в частности, согласованность организации высокочастотных элементов с оптикой глаза и строением сетчатки. Из этой согласованности следует, что любые нарушения в зрительной системе детей в раннем возрасте должны приводить к неправильному формированию высокочастотных элементов. Такое неправильное формирование наблюдается, например, при наличии

у пациентов амблиопии. По данным Флома и др. [Flom et al., 1963a] у амблиопов, как и у наблюдателей с нормальным зрением, тормозные зоны краудинг-эффекта при использовании колец Ландольта и окружающих полос имеют размер, примерно равный размеру минимально видимого стимула. Правда, в настоящее время подобные данные подвергаются ревизии. В сборнике статей Ридберг [Rydberg, 1998] приводятся ссылки на данные, из которых следует, что у амблиопов острота зрения, измеренная по таблицам, меньше остроты, измеренной с помощью изолированных тестов. Более того, отношение величин остроты зрения, определенных с помощью изолированных и сгруппированных символов, в ряде процитированных в сборнике работ предлагается брать за критерий для диагностики амблиопии. Возможно, что такой критерий будет давать удовлетворительный результат при сопоставлении остроты, определенной с помощью решеток и символьный изображений, поскольку эти тестовые изображения могут обрабатываться различными пространственными элементами. Подобное предположение нуждается в дальнейшей проверке и исследовании.

Важными с точки зрения описания механизма краудинг-эффекта являются данные, свидетельствующие о равенстве тормозных зон эффекта у детей 9–12 лет и у взрослых. Сопоставление этих результатов с известными морфологическими и психофизиологическими данными позволяет сделать вывод о локализации механизма краудинг-эффекта именно на уровне стриарной коры, что до настоящего времени не было доказано, а высказывалось только как предположение.

5. Заключение

Предпринята попытка теоретического обоснования полученных данных по измерению остроты зрения и краудинг-эффекта, приняв во внимание физические характеристики изображений и нейрофизиологические механизмы, обеспечивающие описание изображений. Для этого сначала были посчитаны спектры изображений, используемых в экспериментах. Затем оценки остроты, полученные экспериментально с помощью прямоугольных решеток и колец Ландольта, были приведены между собой в соответствие посредством выбора в качестве единицы измерения не величины минимально видимого

разрыва в изображениях, а периода самой высокой пространственной частоты, амплитуду которой в тестовом стимуле может оценивать зрительная система наблюдателя. Таким образом, было показано, что оценки остроты зрения при использовании разных тестовых изображений в задаче их опознания можно согласовать посредством учета спектральных характеристик изображений.

Однако при рассмотрении процесса восприятия нельзя ограничиваться учетом только физических свойств изображений. Спектры изображений должны определяться какими-то структурами, существующими в зрительной системе. Рассмотрение задачи краудинг-эффекта выявило сложность данной проблемы, поскольку показало, что самые высокие частоты в изображениях, от которых зависит определение их формы на пределе разрешения зрительной системы, могут обрабатываться различными пространственными элементами. В частности, такими элементами могут являться «детекторы» полос, краев или «детекторы» решеток. Совокупность всех пространственных элементов обеспечивает пространственно-частотную характеристику зрительной системы, которую естественно рассматривать как суммарный ответ всех элементов на синусоидальные решетки. Мы проводили наши исследования при достаточно высокой средней яркости. В этих условиях получили согласование оценок остроты зрения на разных оптотипах при учете их спектральных характеристик. Известно, что при низких яркостях пространственно-частотная характеристика зрительной системы становится более низкочастотной [см. например, Шелепин и др., 1985, 1992]. Снижение чувствительности к высоким частотам может возникать за счет увеличения размеров зон у пространственных элементов в условиях уменьшения средней яркости, что наблюдается в психофизических экспериментах по изучению пространственных элементов [Blommaert et al., 1981]. Необязательно при этом должно происходить согласованное изменение размеров зон, например, у «детекторов» решеток и полос, поскольку в настоящий момент не имеется прямых доказательств того, что «детекторы» решеток образованы именно «детекторами» полос (простыми РП) по схеме Хьюбела и Визела [Hubel, Wiesel, 1962] как сложные РП. «Детекторы» решеток могут быть организованы и непосредственно РП НКТ. В зависи-

мости от своей организации разные элементы могут быть более или менее чувствительными к разным стимулам в различных условиях наблюдения. Выявленное при низких яркостях противоположное смещение в оценках остроты зрения, полученных при использовании в качестве оптотипов колец Ландольта и прямоугольных решеток [Riggs, 1965], может быть следствием различной изменчивости и чувствительности элементов. (Напомним, что при низких яркостях острота зрения, измеренная как величина минимально видимого разрыва с помощью прямоугольных решеток, наоборот выше, чем острота, измеренная с помощью колец Ландольта.) Таким образом, попытка разрешить проблему расхождения в оценках остроты зрения посредством описания стимулов пространственными элементами выявила ее трудность, заключающуюся в многообразии существующих элементов в зрительной системе и изменчивости этих элементов при смене, например, яркости или времени предъявления стимулов и т. п. Для ряда стимулов в зрительной системе существуют специализированные, более или менее согласованные с ними, пространственные элементы. Эти элементы могут выполнять данную конкретную задачу в наиболее оптимальных для них условиях. При изменении условий наблюдения более чувствительными к стимулу могут уже стать другие элементы. С этой точки зрения решетки и кольца Ландольта могут в разных условиях активизировать разные элементы, что может приводить к рассогласованию оценок.

Видение объектов на пределе разрешения зрительной системы, безусловно, обладает рядом особенностей. Мы подробно рассмотрели феномен краудинг-эффекта и выявили, что краудинг-эффект зависит от многих параметров: от формы тестового стимула, расстояния до окружающих объектов, от пространственно-частотного состава дистракторов и их расположения в пространстве. Мы показали, что при использовании тестовых колец Ландольта и симметричного окружения, тормозное взаимодействие может быть объяснено посредством функционирования самых высокочастотных пространственных элементов — «детекторов» полос и краев. Для объяснения феномена краудинг-эффекта в этом случае нет необходимости вводить в рассмотрение тормозное взаимодействие между соседними элементами. Большая же разница между размерами тормозных зон,

полученная при использовании тестовых колец Ландольта и прямоугольных решеток в случае их окружения решетками со случайной ориентацией, делает затруднительным вывод о том, что феномен краудинг-эффекта во всех случаях является проявлением функционирования только таких элементов. Возможно, что при использовании в качестве тестовых стимулов решеток функционируют другие элементы — «детекторы» решеток. Но имеет под собой основание и другая интерпретация этих результатов. Полученная нами зависимость опознания ориентации решеток на пределе разрешения зрительной системы от используемых дистракторов подобна зависимости, отражающей взаимодействие между пространственно-частотными каналам, выявленной Полатом и Саги [Polat, Sagi, 1993] для больших размеров стимулов. Поэтому в настоящее время мы не можем утверждать, что полученное ухудшение в восприятии решеток при их окружении решетками связано только с функционированием одиночных элементов — «детекторов» решеток. В этом случае, возможно, проявляется и пространственное латеральное взаимодействие между элементами. Необходимо отметить, что Зенгер и Саги [Zenger, Sagi, 1997] показали, что тормозное латеральное взаимодействие между частотными каналами не зависит от фазы маскировочных сигналов, что подтверждает наше предположение об игнорировании фазы «детекторами» решеток.

Для стимулов, подобных буквам, краудинг-эффект, скорее всего, связан с функционированием самых высокочастотных элементов — «детекторов» полос и краев. Об этом может свидетельствовать тот факт, что в большинстве случаев при нормальном зрении у взрослых наблюдателей острота зрения, измеренная по таблицам, не отличается от таковой при измерении ее на изолированных стимулах [Rydberg, 1998]. Расстояния же между символами в таблицах равны их размеру. Кроме того, для описания изображений типа букв важны их фазовые характеристики [Campbell, 1980], которые не могут измерять «детекторы» решеток, нечувствительные к фазе. Подтверждением этого положения могут являться и полученные нами данные по измерению краудинг-эффекта при применении в качестве тестовых стимулов букв Е: тормозные зоны в этом случае оказались аналогичными по размеру зонам, определенным при использовании

тестовых колец Ландольта. Таким образом, для оформленных стимулов наблюдаемое ухудшение в восприятии при близком расположении окружения связано с описанием изображения, состоящего из тестового стимула и дистракторов, высокочастотными элементами типа «детекторов» полос и краев. При таком описании дистракторы попадают в тормозную зону элементов, уменьшая ответ элементов, что и вызывает ухудшение восприятия тестового стимула. В то же время для тестовых решеток, окруженных аналогичными изображениями со случайной ориентацией, ухудшение может быть вызвано тормозным взаимодействием между частотными каналами или функционированием высокочастотных «детекторов» решеток, имеющих большие размеры, чем высокочастотные «детекторы» краев и полос.

Гипотезу о том, что краудинг-эффект может являться следствием взаимодействия между частотными каналами, рассматривали и другие исследователи, занимавшиеся его изучением не на пределе разрешения зрительной системы. В исследовании Чанг и др. [Chung, et al., 2001] при эксцентриситете в 5 угл.град. были взяты стимулы — строки, состоящие из трех букв латинского алфавита. Буквы были в 2 раза большие по размерам, чем предельно видимые стимулы для каждого наблюдателя в этих условиях предъявления. Они были отфильтрованы фильтрами пространственных частот с шириной пропускания в 1 октаву. Две окружающие буквы по отношению к центральной тестовой находились на расстояниях, варьируемых от 0.5 до 3 высот букв. Было показано, что при частотной фильтрации различных букв большую маскировку оказывают буквы с тем же частотным составом, но не выполняется принцип инвариантности по отношению к расстоянию. Напомним, что в исследованиях [Ejima, Tahakashi, 1984; 1985; Polat, Sagi, 1993; 1994] было показано, что тормозное взаимодействие между элементами Габора или несколькими периодами синусоидальных решеток наблюдается на расстоянии 2–3 периодов пространственной частоты, а при больших расстояниях, до 12 периодов, наблюдается фасилитация. Это справедливо во всем диапазоне пространственных частот, использованных в экспериментах (от 0.33 до 12 циклов на градус). То есть в исследовании Чанг и др. не было получено, что, например, при оставлении

в изображениях частот в 2 раза более низких, тормозная зона увеличивается по размеру в 2 раза. Отметим тут также, что при данных условиях наблюдения (при довольно-таки больших размерах стимулов и расстояниях) эти исследователи не получили никакого краудинг-эффекта при предъявлении стимулов в центре поля зрения (минимальное расстояние составляло половину высоты буквы).

В работе Пели и др. [Pelli et al., 2004] при фиксированном расстоянии до отфильтрованных изображений изучался вопрос о влиянии контраста маскировки на опознание центрального стимула. При взаимодействии между частотными каналами [Zenger, Sagi, 1997] маскировочная функция имеет относительно контраста экспоненциальную зависимость. Полученная исследователями зависимость влияния маскировки от контраста носила другой характер. Авторы также фильтровали буквы фильтрами пространственных частот. Предъявление стимулов было при эксцентриситете 4°, буквы имели размер 0.25°. Исследователи показали, что контрастный порог опознания повышается при использовании цветного шума в области 2.5 угловых градусов, то есть в области в 10 раз превосходящей область, занимаемую сигналом. В другом эксперименте на эти же буквы накладывался белый шум. Шум оказывал воздействие только в области, где находился стимул. Пели с соавторами предполагают, что маскировка белым шумом связана с выделением отдельных деталей объекта, в то время как краудинг-эффект не является стандартной формой маскировки сигнала.

В связи с подобием полученных результатов, видимо, следует упомянуть еще работы по выравниванию контраста синусоидальных решеток, расположенных в центральном круге, окруженных кольцами, заполненными другой пространственной частотой с фиксированным контрастом. Канноном и Фулленкампом [Cannon, Fullenkamp, 1991] было показано, что тормозная зона увеличивается в зависимости от контраста окружения. Зона больше по размерам при окружении той же самой пространственной частотой. Максимальный полученный в этих исследованиях размер тормозной зоны был равен 10 периодам центральной пространственной частоты.

В перечисленных выше исследованиях авторы приходят к выводу, что частотный состав изображений влияет на краудинг-эффект,

но сам феномен не является проявлением только взаимодействия между пространственно-частотными каналами. В настоящее время большинство исследователей не предлагают пока какую-либо гипотезу для объяснения краудинг-эффекта для изображений больших размеров. Только Леви и др. [Levi et al., 2002] пришли к заключению, что в фовеа для ряда синтезированных изображений (буквы Е и кольца Ландольта, образованные элементами Габора) краудинг-эффект подобен обычной латеральной маскировке. С ними не согласны Пели с коллегами [Pelli et al., 2004] и другие исследователи. В известной наиболее часто цитируемой работе Страсбургера и др. [Strasburger et al., 1991] показано отсутствие краудинг-эффекта для черно-белых изображений больших размеров при предъявлении их в центре поля зрения. Эти исследователи использовали в качестве тестовых изображений и окружения цифры, предъявляемые на 0.1с. В настоящее время было получено уточнение этим данным: в пределах оптической функции рассеяния краудинг-эффект все же наблюдается для изображений больших размеров. Однако для других типов изображений механизм краудинг-эффекта описать не удается [Ehrt and Hess, 2005]. Учитывая приведенные данные, вывод можно сформулировать следующим образом: оформленные изображения и решетки могут при определенных условиях предъявления стимулов обрабатываться разными пространственными элементами, существующими в зрительной системе, поэтому полного соответствия между их взаимодействием не наблюдается. Для оформленных изображений на пределе разрешения зрительной системы ухудшение опознания тестового стимула объясняется попаданием дистракторов в зоны описывающих стимул пространственных элементов, подобных «детекторам» полос и краев. Для текстурированных стимулов, например, для решеток, ухудшение их опознания может быть связано как с функционированием других элементов — «детекторов» решеток, так и с взаимодействием между элементами. Механизм же краудинг-эффекта для изображений большого размера нуждается в дальнейшем изучении.

Подтверждением наших предположений о возможности описания изображений различными элементами могут служить также следующие данные. В офтальмологической практике описаны ред-

кие случаи присутствия у пациентов нормальной остроты зрения, определенной с помощью решеток, и сниженной при использовании символьных или буквенных таблиц [Хювяйринен, 1996]. При этом пациенты не страдали зрительной агнозией. С нашей точки зрения, такое явление вполне может иметь место, если учитывать, что решетки и оформленные изображения могут описываться различными пространственными элементами.

Список литературы

Бондарко В. М., Данилова М. В. Краудинг-эффект при различном окружении и расстоянии между стимулами // Сенсорные системы. 1996. Т. 10. № 1. С. 132–140.

Бондарко В. М., Данилова М. В. Влияние окружающих стимулов на определение ориентации прямоугольных решеток на пределе разрешения зрительной системы человека // Физиология человека. 1998. Т. 24. N5. С. 65–71.

Бондарко В. М., Данилова М. В. Зависимость размера тормозной зоны в краудинг-эффекте от формы тестового стимула // Оптический журнал. 1999. Т. 66. N10. С. 26–30.

Бондарко В. М., Данилова М. В., Красильников Н. Н., Леушина Л. И., Невская А. А., Шелепин Ю. Е. Пространственное зрение. Спб.: Наука. 1999. 218 С.

Бондарко В. М., Данилова М. В. Связь краудинг-эффекта с функционированием высокочастотных пространственных элементов // Сенсорные системы. 2002. Т. 16. N2. С. 89–99.

Бондарко В. М., Семенов Л. А. острота зрения и краудинг-эффект у школьников 8–17 лет // Физиология человека. 2005. Т. 31. N5. С. 13–21.

Бондарко В. М., Семенов Л. А. Острота и гиперострота зрения у школьников 11–17 лет // Физиология человека. 2012. Т. 38. № 3. С. 56–61.

Волков В. В., Колесникова Л. Н., Шелепин Ю. Е. Частотно-контрастные характеристики и острота зрения в офтальмоло-

гической практике // Офтальмол. журн.. 1983. N3.С. 142–151.

Данилова М. В., Бондарко В. М. Влияние контекста на выполнение задачи различения ориентации прямоугольных решеток на пределе разрешения зрительной системы // Сенсорные системы. 2002. Т. 16. N2. С. 100–109.

Кравков С. В. Глаз и его работа. М. — Л.: Изд. Академии наук СССР. 1950. 532 с.

Невская А. А., Леушина Л. И., Бондарко В. М. Исследование формирования зрительных представлений об окружающем мире у детей первого года жизни //Физиология человека. 1996. Т. 22. № 3. С. 57–63.

Невская А. А., Леушина Л. И., Бондарко В. М. Формирование наглядно-образного мышления у младенцев первого года жизни с дефектами зрения //Физиология человека. 1998. Т. 24(3). С. 20–26.

Семенов Л. А., Чернова Н. Д., Бондарко В. М. Измерение остроты зрения и краудинг-эффекта у детей в возрасте от трех до девяти лет // Физиология человека. 2000. Т. 26. С. 67–72.

Рожкова Г. И., Токарева В. С., Ващенко Д. И., Васильева Н. Н. Возрастная динамика остроты зрения у школьников. Бинокулярная острота зрения для дали // Сенсорные системы, 2001, Т. 15, № 1, с. 47–52.

Фарбер Д. А., Семенова Л. К., Алферова В. В. и др. Структурно-функциональная организация развивающегося мозга. Л.: Наука.1990. 198с.

Хювяринен Л. Зрение у детей: нормальное и с нарушениями. СПб.: Петербург -XXI век, 1996. 72с.

Шамшинова А. М., Волков В. В. Функциональные методы исследования в офтальмологии. М.: Медицина. 1998. 416 С.

Шелепин Ю. Е., Бондарко В. М. Разрешающая способность и дискретизация изображений в зрительной системе // Рос. фи-

зиологический ж. им. И. М. Сеченова. 2002. Т. 88. N9.С. 1116–1132.

Шелепин Ю. Е., Глезер В. Д., Бондарко В. М., Павловская М. Б., Вол И. А., Данилов Ю. П. Пространственное зрение // Физиология зрения / Ред. А. Л. Бызов. М.: Наука, 1992. С. 528–586.

Шелепин Ю. Е., Колесникова Л. Н., Левкович Ю. И. Визоконтрастометрия. Измерение пространственных передаточных функций зрительной системы. Л.: Наука, 1985.

Шелепин Ю. Е., Волков В. В., Макулов В. Б., Колесникова Л. Н., Паук В. Н. Измерение функциональных возможностей зрительной системы человека // Вестн. АН СССР. 1987. N9. С. 63–72.

Atkinson J. Review of human visual development: crowding and dyslexia // Vision and visual dysfunction (ed.J.R.Cronly- Dillon). 1991. V. 13. Vision and visual dislexia (ed.J.F.Stein). P. 44–57.

Atkinson J., Braddick O., Braddick F. Acuity and contrast sensitivity of infant vision // Nature. 1974. Vol. 247. № 5440. P. 403–404.

Atkinson J., Braddick O., Moar K. Development of contrast sensitivity over the first 3 months of life in the human infant. // Vision Res.1977. Vol.17. P. 1037–1044.

Banks M. S., Bennett P. J. Optical and photoreceptor immaturities limit the spatial and chromatic vision of human neonates. // J. Opt. Soc. Amer. 1988. Vol. 5. P. 2059–2079.

Banks M. S., Salapatek P. Infant pattern vision: A new approach based on the contrast sensitivity function // J. Exper. Child Psychol. 1981. Vol.31. P. 1–45.

Banks M. S., Stephens B. R., Hartman E. E. The development of basic mechanisms of pattern vision: Spatial frequency channels. // J. Exper. Child Psychol. 1985. Vol. 40. P. 501–527.

Blakemore C., Campbell F. W. On the existence in human visual

system of neurones selectively sensitive to the orientation and size of retinal image // J. Physiology. 1969. V. 203. N1. P. 237–260.

Blommaert F.J.J., Heynen H.G.M., Roufs J.A.J. Point spread function, analysis and use // IPO Annu. Prog. Rep. 1981. Vol. 16. P. 61–79.

Bondarko V.M., Danilova M.V. What spatial frequency do we use to detect the orientation of a Landolt C? // Vision Res. 1997. Vol. 37. P. 2153–2156.

Bonneh Y., Sagi D. Effects of spatial configuration on contrast detection // Vision Res. 1998. Vol. 38. P. 3541–3553.

Bonneh Y., Sagi D. Configuration saliency revealed by dichoptic masking // Vision Res. 1999. Vol. 39. P. 271–281.

Bouma H. Interaction effects in parafoveal letter recognition // Nature. 1970. Vol. 226. № 5241. P. 177–178.

Campbell F.W. The physics of visual perception // Phil. Trans. R. Soc. Lond. Ser.B. 1980. Vol. 290. № 1038. P. 5–9.

Campbell F.W., Robson J.G. Application of Fourier Analyses to the Visibility of Gratings // J. Physiol. 1968. V. 197. P. 551–566.

Campbell F.W., Gubisch R.W. Optical quality of the human eye // J. Physiol. 1966. Vol. 186. № 3. P. 558–578.

Campbell F.W., Shelepin Y.E. The mechanics of the foveola and its role in defining an object // Perception. 1989. Vol. 18. № 4. P. 532.

Campbell F.W., Shelepin Y.E., Tegeder R.W., Pavlov N.N. Psychophysical measurement of the intercone separation and object recognition in the human foveola // Ophthal. Physiol. Opt. 1992. Vol. 12. N1. P. 101–102.

Cannon M.W., Fullenkamp S.C. Spatial interactions in apparent contrast: inhibitory effect among grating patterns of different spatial frequencies, spatial positions and orientations // Vision Res. 1991. Vol. 31. № 11. P. 1985–1998.

Danilova M. V., Bondarko V. M. Foveal contour interactions and crowding effects at the resolution limit of the visual system // J. Vision. 2007. V. 7(2). N. 25. P. 1–18.

Dobson V., Teller D. Y. Visual acuity in human infants: A review and comparison of behavioural and electrophysiological studies // Vision Research. 1978. Vol. 18. P. 1469–1483.

Ehlers H. Clinical testing of visual acuity // A.M.A. Archives of Ophthalmology. 1953. Vol. 49. P. 431–434.

Ehrt O., Hess R. Foveal contour interaction: detection and discrimination // J. Opt. Soc. Am. A. 2005. Vol. 22(2). P. 209–216.

Ejima Y., Takahashi S. Facilitatory and inhibitory after-effect of spatially localized grating adaptation // Vision Res. 1984. Vol. 24. N9. P. 979–985.

Ejima Y., Takahashi S. Effect of localized grating adaptation as a function of separation along the length axis between test and adaptation areas // Vision Res. 1985. Vol. 25. 1701–1707.

Fantz R. L. The origin of form perception // Scient. Amer. 1961. Vol. 204. № 5. P. 66–72.

Fantz R. L., Ordy J. M., Udelf M. Maturation of pattern vision in infants during the first six months // J. Comp. Physiol. Psychol. 1962. V.55. P. 907–917.

Field D. J. Relations between the statistics of natural images and the response properties of cortical cells // J. Opt. Soc. Am. A. 1987. Vol. 4. P. 2379–2394.

Fiorentini A., Pirchio M., Spinelli D. Electrophysiological evidence for spatial frequency selective mechanisms in adults and infants. // Vision Res.1983. Vol. 23. P. 119–127.

Flom M. C. Contour interaction and the crowding effect // In R. P. Rutstein (Ed.). Problems in Optometry. 1991. V. 3. № 2. P. 237–257.

Flom M. C., Weymouth F. W., Kahneman D. Visual resolution and contour interaction // J. Opt. Soc. Am. 1963a. V. 53. P. 1026–1032.

Flom M. C., Heath G., Takahaski E. Contour interaction and visual resolution: Contralateral effect // Science. 1963b. V. 142. P. 979–980.

Hess R. F., Dakin S. C., Kapoor N. The foveal contour interaction: physics or physiology? // Vision Res. 2000. Vol. 40. P. 365–370.

Hubel D. H., Wiesel T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex // J. Physiol. 1962. V. 160. P. 251–260.

Huttenlocher P. R., DCourten C., Garei L. J., Wan der Loos H. Synapsogenesis in human visual cortex — evidence for synapse elimination during normal development // Neurosci. Letters. 1982. Vol. 33. P. 247–252.

Hyvarinen L., Colenbrander A., Mayer L., Faye E. E., Hood C. M., Bayley I., Ardity A., Lang M. A., Seidman K., Kopidlansky E., Greening C. The LH Symbol tests. A system for vision testing in children // N. — Y.: Lighthouse Low Vision Products. A Division of the Lighthouse Inc. 1992.

Landolt E. Opto-types simples. Paris. O. Doin. 1889. 112 p.

Lindstedt E. Manual BUST-D, a vision test for young child and/or handicapped child / adult. Stockholm, Sweden: Elysin. 1988.

Levi D. M. Crowding — An essential bottleneck for object recognition: A mini-review //Vision research. 2008. V. 48. P. 635–654.

Mayer T. Experiments on visual acuity. // Spatial Vision. 1755/1987. Vol. 2. P. 81–97.

Morgan M. J. Hyperacuity // In: Spatial vision. Regan D. (Ed.) 1991. London: Macmillan. P. 87–113.

Nazir T. A. Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision // Vision Res. 1992. Vol. 32. № 4. P. 771–777.

Norcia A. M., Tyler C. W., Hamer R. D. Development of contrast sensitivity in the human infant. // Vision Res. 1990. Vol.30. P. 1475–1486.

Pelli D. G., Palomares M., Majaj N. J. Crowding is unlike ordinary masking: distinguishing feature detection and integration // J. Vision. 2004. V. 4. P. 1136–1169.

Polat U., Sagi D. Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiment // Vision Research. 1993. V. 23. № 7. P. 993–999.

Polyak S. L. The retina. Chicago: University of Chicago Press. 1941. 607 p.

Reuter J. B. Kent Infant Development Scale Manual. 2d ed. // Kent (Ohio): Kent-Development Matrix Inc. 1985.

Riggs L. A. Visual acuity // In: Vision and visual perception. 1965. New York: Wiley. P. 321–349.

Rydberg A. Clinical assessment of visual function. With particular emphasis on testing methods for young children. Stockholm. 1998.

Sheridan M. D. New appliances, Sheridan Gardiner test for visual acuity. // Brit. Medical J., 1970. Vol. 108, p.9–15.

Shlaer S. The relation between visual acuity and illumination // J. gen. Physiol. 1937. Vol. 21. P. 165–188.

Snellen E. Echelle typographique pour mesurer l’acuite de la vision. 1862.Utrecht.

Strasburger H., Harvey L. O., Jr., Rentschler I. Contrast thresholds for identification of numeric characters in direct and eccentric view // Perception and Psychophysics. 1991. Vol. 49. № 6. P. 495–508.

Stuart J. A., Burian H. M. A study of separation difficulty. Its relationship to visual acuity in normal and ambliopic eyes // Am. J. Ophthalmology. 1962. Vol. 53. P. 471.

Teller D. Y. The forced-choice preferential looking procedure: A psychophysical technique for use with human infants // Infant Behav. Dev. 1979. V.2. P. 135–153.

Teller D. Y. First glances: The vision of infants. The Friedenwald Lecture // Invest. Ophthalm. Visual Sci. 1997. V.38. № 11. P. 2183–2203.

Teller D. Y., Morse R., Borton R., Regal D. Visual acuity for vertical and diagonal gratings in human infants. // Vision Res. 1974. Vol.14. P. 1433–1439.

Townsend J. T., Taylor S. G., Brown D. R. Lateral masking for letters with unlimited viewing time // Perception and Psychophysics. 1971. Vol. 10. P. 375–378.

Westheimer G. Visual hyperacuity // Progress in Sensory Physiology. 1981. Vol. 1. P. 1–30.

Wilson H. R. Development of spatiotemporal mechanisms in infant vision. // Vision Res. 1988. V.28. P. 611–628.

Wilson H. R., Gelb D. J. Modified line element theory for spatial frequency and width discrimination // J. Opt. Soc. Amer. A. 1984. V. 1. 124–131.

Wolford G., Chambers L. Lateral masking as a function of spacing // Perception and Psychophysics. 1983. Vol. 33. P. 83–96.

Woodrow H. The effect of pattern upon simultaneous letter span. // Am. J. Psychology. 1938. Vol. 51. P. 83–96.

Yuodelis C., Hendrickson A. A. qualitative and quantitative analysis of the human fovea during development // Vision Res. 1986. Vol.26. P. 847–855.

Zenger B., Sagi D. Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection // Vision Res. 1996. Vol. 36. P. 2497–2513.

Крупномасштабные нейронные сети головного мозга человека

Глава 4.
Эффект МакКоллаф — ключ к пониманию особенностей зрительного анализа

А. Р. Кезели

Центр экспериментальной биомедицины им. И. С. Бериташвили.
Тбилиси, Грузия

Обработка информации, поступающей с сетчатки глаза, конечным результатом чего является возникновение зрительного образа, несомненно принадлежит к числу сложнейших задач, решаемых в нейронных сетях головного мозга человека и большинства животных. Формирование зрительного образа невероятно сложный, многоступенчатый процесс, протекающий при параллельном участии различных подкорковых структур и корковых областей зрительного анализатора.

Как известно, нейронные структуры зрительной системы человека и многих высших позвоночных к моменту рождения уже окончательно сформированы и готовы к функционированию, но нейронные сети еще не загружены. Для того, чтобы начать функционировать, система должна «научиться видеть». Процесс обучения нейронной сети зрительного анализатора начинается в первый же момент поступления информации на вход сети и, по всей вероятности, продолжается всю жизнь. Зрительный образ окружающего нас мира возникает на основе обработки поступающих с глаза огромных массивов информации, изначально содержащей в себе все данные, характеризующие воспринимаемый зрением объект. Опираясь на эту информацию, зрительная система должна научиться создавать зрительный образ и фиксировать его в памяти. «Все, что увидено, должно быть запомнено!» — основной принцип действия нейронной сети зрительного анализатора. Только в таком случае зрение, как высшая функция, приобретает биологический смысл. Довольно часто информация, поступающая со входа, может быть неполной, может требовать коррекции, или быть неоднозначной, допускающей различные интерпретации. Исходя из этого, следует допустить, что должны существовать некие гипотетические механизмы, обеспечи-

вающие в таких случаях стабильное и адекватное опознание объектов окружающего мира. Это должны быть механизмы компенсации и коррекции, интерпретационный, и некоторые другие механизмы. Исследование этих механизмов сопряжено с большими трудностями, поскольку они практически недоступны для прямых экспериментальных подходов. Тем не менее, достаточно значимые данные можно получить, изучая косвенные результаты их функционирования. В этом отношении одним из плодотворных подходов может оказаться изучение зрительных иллюзий, возникновение которых в большинстве случаев связано с активацией именно этих механизмов, результаты действия которых можно в этом случае" видеть своими глазами" в буквальном смысле этого слова.

Среди механизмов, обеспечивающих стабильность зрительного восприятия, особое место принадлежит механизмам константности цветовосприятия, назначение которых — обеспечить неизменность воспринимаемого цвета поверхности при довольно широких изменениях спектра источника освещения. Исследуя механизмы константности [Джафарли и др.,1991; Jafarli et al., 1993], мы обратили внимание на особый класс зрительных эффектов, которые принципиально отличаются от зрительных иллюзий — это так называемые обусловленные последовательные эффекты и, в том числе, самый известный и, вероятно, самый изучаемый из них — т. н. эффект МакКоллаф (ЭМК). Впервые описанный в 1965 году Челестой МакКоллаф [McCollough, 1965], этот эффект на протяжении более полувека остается предметом особого внимания исследователей. Суть его состоит в том, что если последовательно предъявлять наблюдателю черно-цветные пространственные решетки противоположной ориентации и дополнительных цветов (например, черно-красную горизонтальную и черно-зеленую вертикальную) (рис. 1), с темным интервалом между предъявлениями, и после примерно 20–25 сочетаний предъявить черно-белую решетку такой же, или близкой к ней пространственной частоты (рис. 2), то белые полосы тестовой решетки кажутся наблюдателю окрашенными в дополнительные цвета (в нашем случае горизонтальные линии будут казаться зеленоватыми, а вертикальные-красноватыми). Пристальный интерес к ЭМК вызван следующими особенностями:

1. В отличие от зрительных иллюзий, для его возникновения требуется выработка — эффект возникает только после многократного предъявления адаптационных стимулов.
2. Большая продолжительность его сохранения — после выработки эффект сохраняется часами, днями и даже неделями.

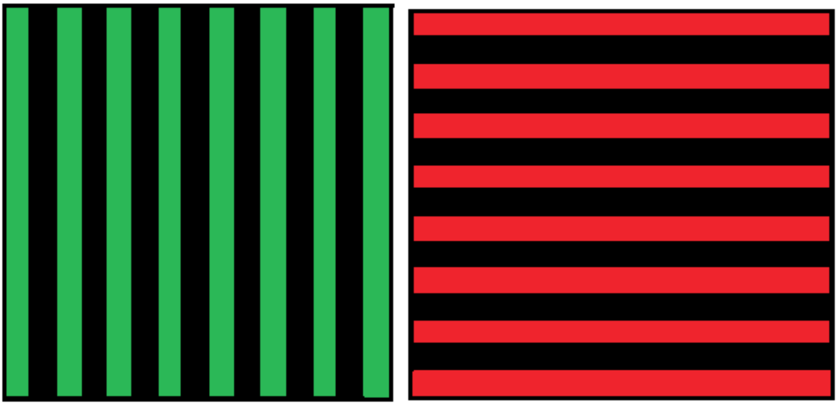


Рис. 1. Адаптационная решетка

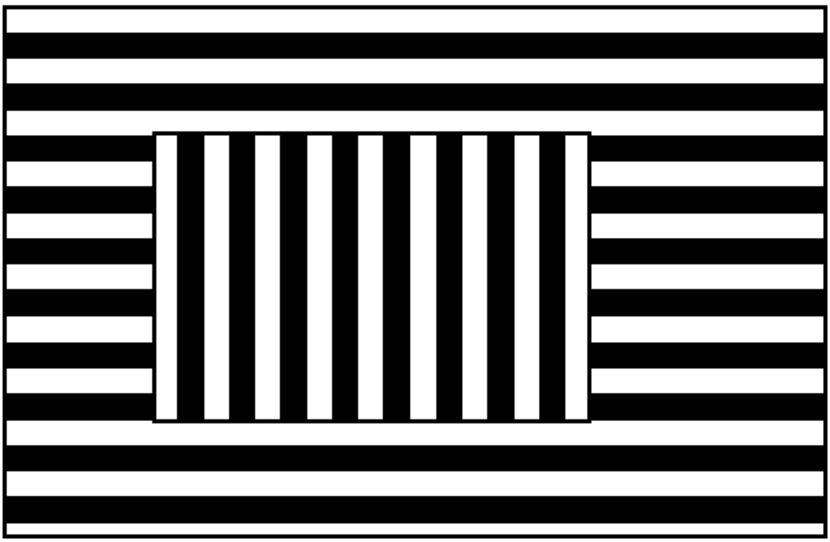


Рис. 2. Тестовая решетка

Естественно, что сразу возник вопрос о механизмах его возникновения. За прошедшие более чем 50 лет эффект исследован подробнейшим образом, созданы десятки его модификаций, уточнены многие детали и выявлены особенности его возникновения, опубликованы сотни статей, предложены различные гипотезы, от нейронно — детекторной вплоть до Павловского условно-рефлекторного обусловливания [Bedford et al., 1993; Barnes et al., 1999; Allan et al., 1991; Dodwell, Peter et al., 1990; Favreau et al., 1972; Jones, Paul et al., 1975; McCollough, 2000; Murch, Hirsch, 1972; Sheth, Shimojo, 2008; Siegel, S., Allan, 1987; Stromeyer, 1969; Stromeyer, 1978; White et al., 1978][1]. Однако ни одна из них не согласуется полностью с множеством экспериментальных данных, накопленных в мировой литературе так что по сей день вопрос остается открытым.

На основе результатов собственных многолетних исследований ЭМК [Кезели и др., 2001; Intskirveli et al., 2000; Malania et al., 2002; Sharikadze, Kezeli, 2002], нами была выдвинута гипотеза о том, что выработка эффекта может быть связана с активацией механизмов константности цветовосприятия [Кезели и др., 1991; Malania et al., 2004]. Дело в том, что адаптационные стимулы — черно-цветные решетки, по своей сути неоднозначны, поскольку черно-цветная решетка может быть генерирована в абсолютно различных ситуациях — это может быть, например, черно-красная решетка, освещенная белым светом, или черно-белая решетка, освещенная красным. Если в поле зрения нет никаких указателей, позволяющих определить источник освещения, то ни глаз человека, ни любой физический прибор не смогут различить эти две ситуации, так как красная поверхность при белом освещении отражает физически такой же состав света, что и белая поверхность при красном освещении. А так как в процессе выработки эффекта информации об источнике света нет, при смене предъявляемых стимулов система должна определить — то ли красно-черная решетка заменяется черно-зеленой при неизменном освещении, то ли это одна и та же черно-белая решетка, освещаемая то красным, то зеленым светом. В случае с возможным изменением спектра источника

[1] Как уже было сказано, за эти годы накопилось огромное количество публикаций, посвященных этой теме, поэтому в списке литературы мы приводим лишь малую, основную часть этих работ.

освещения, должны активироваться механизмы константности цветовосприятия, назначение которых — выработать соответствующие поправки на освещение. Для этого в первую очередь система должна получить информацию об источнике. В естественных условиях спектр источника можно определить, если в поле зрения находятся белые или зеркальные поверхности, полностью отражающие падающий на них свет, или же цветные поверхности, селективно отражающие в различных областях спектра [Максимов, 1984]. В искусственных условиях выработки ЭМК единственным стабильным показателем, однозначно связанным с возможным цветом источника освещения, может служить только ориентация решетки (в нашем примере прослеживается такая связь: горизонтальная ориентация — источник света возможно красный, вертикальная — зеленый). Если система установит связь между ориентацией решетки и возможным спектром источника (что и должно происходить в результате неоднократного сочетания ориентации и цвета полос адаптационных решеток), то при предъявлении тестовой решетки механизм константности цветовосприятия должен снизить чувствительность к красному при горизонтальной ориентации (в результате горизонтальная белая полоса покажется зеленоватой), а при вертикальной — к зеленому (белая вертикальная должна будет казаться красноватой). По сути, именно это и происходит при выработке ЭМК. Таковы чисто теоретические предпосылки, для подтверждения (или опровержения) которых нами были проведены десятки экспериментов, в которых мы искали ответы на существующие вопросы.

Методически все серии экспериментов проводились в одинаковых условиях: испытуемый глядел сквозь смотровое окно внутрь темного ящика, на задней стенке которого предъявлялись используемые стимулы (рис.3)

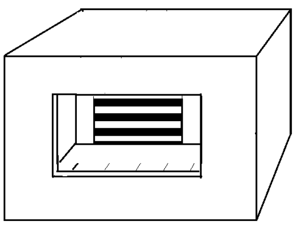


Рис. 3. Экспериментальный ящик

Серия 1:
Неоднозначность предъявляемых стимулов

В этой серии экспериментов мы применяли стимулы, практически идентичные с «классическими» решетками, с одним лишь различием: вместо черно-цветных решеток мы использовали бело-цветные. Такие решетки имеют такие же основные параметры (ориентация, цвет, пространственные частоты), что и черно-цветные, с одной лишь разницей — вместо черных полосок белые (рис.4)

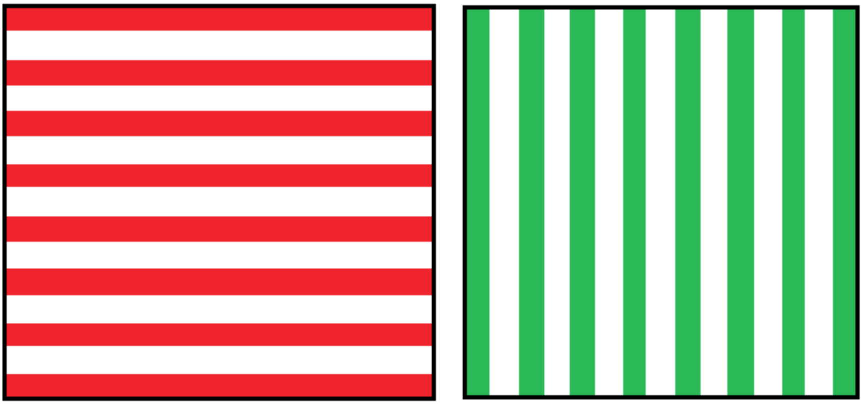


Рис. 4.

Наличие белых полосок исключает неоднозначность решеток. При замене бело-красной на бело-зеленую они указывают на то, что цвет источника освещения не менялся, просто одна решетка заменила другую. Если, согласно нашей гипотезе, выработка ЭМК связана с активацией механизмов константности, вносящих поправку на освещение, в этом случае эффект вырабатываться не должен, поскольку система «знает», что освещение не менялось. И в самом деле, если использовать в качестве адаптационных красно-белую и зелено-белую решетки в стандартных условиях выработки ЭМК, эффект не вырабатывается (эксперименты проведены на 32 испытуемых обоего пола –15 мужского, 17 женского, в возрасте от 15 до 75 лет). Ни у одного из них эффект не выработался. Для контроля на том же контингенте через два дня провели серию с выработкой

обычного ЭМК. У 29 из 32-х, т.е. примерно у 90 %, эффект выработался (из литературы известно, что в норме примерно у 80 % эффект выработать не удается. Кстати, факт, также требующий объяснения).

Судя по этим результатам, можно придти к выводу, что если исключить неоднозначность входящей информации, эффект типа ЭМК вырабатываться не должен.

Серия 2.
Информация о возможном источнике освещения

Целью этой серии являлось, выяснить, как влияет на результат предоставление системе возможности, определить спектр источника. Эти эксперименты состояли из нескольких ступеней. Вначале вырабатывали обычный ЭМК — с последовательным предъявлением черно-цветных решеток при белом освещении. Участвовало 25 испытуемых. 24 четко видели эффект и они были отобраны для дальнейшего участия в эксперименте. Через несколько дней, после того, как тестирование показывало, что ранее выработанный эффект отсутствует, на том же контингенте вырабатывали эффект в новых условиях. Предъявлялись черно-белые решетки противоположной ориентации, освещаемые скрытым от наблюдателя источником то красным, то зеленым (в зависимости от ориентации) светом. Как уже было сказано, глазом различить эти две ситуации невозможно. Система воспринимает эти две различные ситуации одинаково и, как и следовало ожидать, у всех из них эффект выработался. Таким образом мы убедились, что в обеих ситуациях — как при неизменном источнике и черно-цветных стимулах, так и при меняющемся освещении и черно-белых стимулах, эффект вырабатывается одинаково.

Затем, опять выждав время, необходимое для исчезновения выработанного до этого эффекта, мы внесли изменения в дизайн эксперимента — поместили между наблюдателем и адаптационными стимулами прозрачное стекло с наклееными на него маленькими кусочками цветных бумажек. Источники освещения остались прежними, разница была лишь в том, что свет, освещающий стимулы, проходя через стеклянную перегородку, освещал также и цветные бумажки (рис.5).

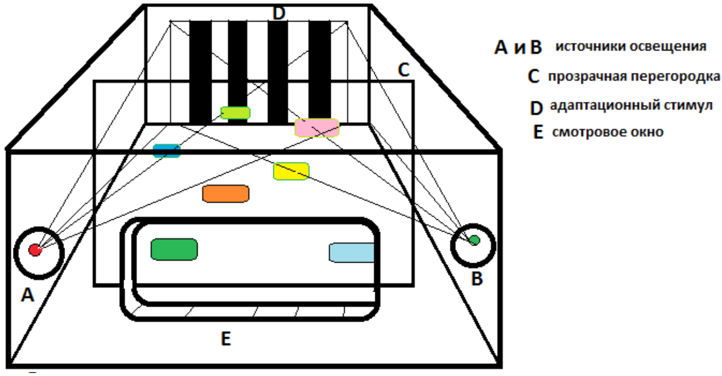


Рис. 5.

Были исследованы две возможные ситуации:

А. Источник света белый, стимулы цветные. Последовательность предъявления стимулов — черно-красный горизонтальный, черно-зеленый вертикальный. Цветные бумажки, находящиеся при этом в поле зрения и освещаемые одновременно со стимулами, указывают на то, что источник света не меняется. Последовательное предъявление черно-цветных стимулов при белом освещении — стандартная ситуация выработки ЭМК и эффект должен вырабатываться, однако в данном случае исключается неоднозначность источника. Если наша гипотеза верна, эффект в этом случае вырабатываться не должен.

И действительно, ни у одного из испытуемых эффект выработать не удалось.

Б. Стимулы черно-белые, цвет источников освещения меняется. Последовательность предъявления стимулов — черно-белый горизонтальный, освещение красное, черно-белый вертикальный, освещение зеленое. Изменение при этом цвета бумажек дает системе возможность, вычислить спектр освещения и внести соответствующие поправки. Если это произойдет, эффект должен выработаться.

У 20-ти из 24 испытуемых наблюдался четко выраженный эффект, двое эффект видели, но слабо, у еще двоих эффект не выработался.

Серия 3.
Раздельное предъявление источника и показателей освещения

В этой серии мы попытались выяснить, возможно ли выработать последовательный эффект, если показатель освещения предъявлять отдельно от предполагаемого источника.

В том же экспериментальном ящике наблюдатель видел через смотровое окно белую трехугольную пирамиду, стоящую внутри ящика перед задней стенкой и освещаемую сверху скрытым от глаз источником (рис.6)

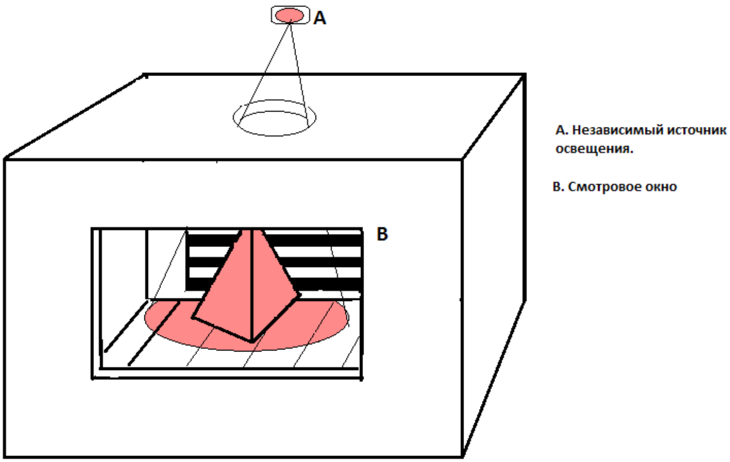


Рис. 6.

Пирамида освещалась то красным, то зеленым светом. Одновременно со сменой источника на задней стенке менялась ориентация предъявляемой черно-белой решетки. Таким образом, устанавливалась закономерность: при предъявлении, скажем, горизонтальной решетки освещение всегда красное, а при вертикальной всегда зеленое. Если система, ищущая показатели освещения, зафиксирует эту закономерность, она должна будет вносить поправки на освещение.

И в самом деле, несмотря на то, что наблюдателю ни разу не предъявлялась окрашенная решетка, при разглядывании тестовой черно-белой решетки 19 из 24-х видели белые полосы окрашенными

в соответствующие цвета, указывающие на то, что и в таких условиях возможно выработать эффект МакКоллаф.

Суммирая данные, полученные как в этом, так и в остальных приведенных наших экспериментах, можно заключить, что при изменении спектра источника система в любой ситуации ищет стабильный признак, однозначно связанный с цветом освещения и вырабатывает соответствующие компенсационные сдвиги.

Обсуждение

Приведенные результаты дают нам все основания утверждать, что наша гипотеза о том, что ЭМК возникает в результате активации механизмов константности цветовосприятия, должна быть верна.

С другой стороны, изучение особенностей этого чрезвычайно интересного феномена, как нам кажется, дает ключ к пониманию более общих проблем деятельности сенсорных систем и, в частности, зрительной системы.

Особенностью сенсорного анализа является тот факт, что ни одна сенсорная система не оперирует абсолютными значениями измеряемых параметров, хотя очень тонко настроена на детекцию изменений. Теплее или холоднее, тяжелее или легче, громче или тише, ярче или темнее — такие сравнения соответствуюшие системы способны определять с невероятной точностью, однако в результате адаптационных сдвигов сильно затруднено определение абсолютных значений измеряемых параметров. В качестве примера можно привести общеизвестный факт восприятия черного цвета: черные буквы печатного текста на солнце гораздо ярче, чем белая бумага в помещении, однако в обеих ситуациях буквы воспринимаются черными, а бумага белой. Этот и многие другие факты указывают на то, что в каждой заданной ситуации система как бы заново «калибруется», определяя нулевой уровень, от которого ведется отсчет. С биологической точки зрения это оправдано, поскольку дает возможность во много раз расширить шкалу значений воспринимаемых раздражителей. В случае зрительной системы результатом является почти абсолютный порог восприятия — от нескольких десятков квантов света и невероятно широкий диапазон, вплоть до яркого солнца.

Учитывая, что в числе ситуаций, встречающихся в естественных условиях, немалое их количество является привычным, повторяющимся, представляется логическим допустить, что опять-таки исходя из биологической целесообразности, выгодно иметь готовые адаптационные сдвиги, выработанные и зафиксированные в случаях часто повторяющихся ситуаций. Судя по всему, именно это и происходит при выработке ЭМК — компенсационные сдвиги, вырабатываемые в неоднозначных ситуациях, фиксируются в повторяющихся ситуациях и при обнаружении признаков этой ситуации, осуществляют соответствующие корректировки.

Тут уместно вспомнить труды грузинского психолога Дмитрия Узнадзе, автора т.н. «теории установки» [Узнадзе, 1949]. Согласно Узнадзе, установка означает неосознаваемую готовность субъекта к восприятию будущих событий и действиям в определенном направлении. Он показал, что установка вырабатывается, если при взаимодействии субъекта со средой возникает потребность и существует ситуация ее удовлетворения. В нашем конкретном случае можно сказать, что процедура выработки ЭМК соответствует условиям возникновения установки: в неоднозначной ситуации возникает потребность определить источник освещения. Удовлетворение этой потребности возможно, нужно лишь определить стабильный показатель освещения. Возникновение ЭМК указывает на то, что такой показатель установлен, и что активировались механизмы константности цветовосприятия, выработавшие в системе цветового зрения компенсационные сдвиги, соответствующие данной конкретной ситуации.

Узнадзе разрабатывал положения теории установки применительно к поведенческим актам субъекта. Если применить их к процессам восприятия, то можно дать такую формулировку: установка сенсорной системы означает ее готовность к адаптированному к данной конкретной ситуации восприятию — установке «нулей» на шкале измерений значений внешних раздражителей.

С этих позиций эффект МакКоллаф можно рассматривать, как экспериментальное подтверждение того, что зрительная система, (также как, вероятно, и другие сенсорные системы) постоянно осуществляет и фиксирует на подсознательном уровне компенсацион-

ные сдвиги, обеспечивающие стабильность восприятия в меняющихся ситуациях.

С другой стороны, можно также заключить, что при выработке эффектов типа МакКоллаф, также как и других обусловленных последовательных эффектов, в процессе многократного повторения, нейронная сеть «обучается» узнавать ситуацию, создает «подпрограмму» для выработки основных компенсационных сдвигов в данной ситуации и сохраняет ее в долгосрочной памяти.

Литература

Джафарли М., Максимов В., Кезели А., Антелидзе Н. 1991. О константности цветовосприятия у обезьян.Сенсорные системы, т. 5, No3, 41–46.

Кезели А. Р., Джанелидзе Д. О., Лебанидзе Т. Дж., Цискаришвили Н. М. Последовательные цветовые эффекты и константность цветовосприятия, Известия АН ГССР, серия биологическая, т. 17, 4, 1991, 221–226.

Кезели А., Джанелидзе Д., Лебанидзе Т., Кезели Т. 2001 Цветовые последовательные эффекты и механизмы компенсации и коррекции. Сенсорные системы, N1, т. 15, 36–43.

Максимов В. В. Трансформация цвета при изменениях освещения. М.: Наука, 1984.

Узнадзе Д. Н. Экспериментальные основы психологии установки. — Москва, 1949.

Allan, L.G., Siegel, S., Toppan, P. 1991 Assessment of the McCollough effect by a shift in the psychometric function. Bulletin of the Psychonomic Society, 29, 21–24.

Bedford, F. L., & Reinke, K. S. 1993. The McCollough effect: Dissociating retinal from spatial coordinates. Perception & Psychophysics, 54, 515–526.

Barnes, J., Howard, R.J., Senior, C., Brammer, M., Bullmore, E.T., Simmons, A., et al. 1999. Brain imaging: The functional anat-

omy of the McCollough contingent color after-effect. NeuroReport, 10, 195–199.

Dodwell, Peter C.; Humphrey, G. Keith 1990. “A functional theory of the McCollough effect”. Psychological Review. 97 1: 78–79.

Favreau O. E., Emerson V. F., Corballis M. C. 1972 Motion perception: a color-contingent aftereffect. Science 176:78–79.

Intskirveli I, Kezeli A., Kezeli T., Lebanidze T., Janelidze D. 2000. Contingent Color Aftereffects and Changes of Color Sensitivity. Bulletin of The Georgian Academy of Sciences. 162, N.1, 141–143.

Intskirveli, I., Kezeli, A., Malania, M., Roinishvili, M. & Janelidze, D. 2002. The Color Aftereffects Occur with Participation of Mechanisms of Color Constancy. Proceedings of the Georgian Academy of Sciences, Biological Series.No.1–2, Vol.28, 65–71.

Jafarli, M., Kezeli, A. & Maximov, V. 1993. Is the primate cortical area V4 important for colour constancy? Proc. of The Georgian Acad. of Sci. Biological series, vol.19, N4, 221–225.

Jones, Paul D.; Holding, Dennis H. 1975. “Extremely long-term persistence of the McCollough effect”. Journal of Experimental Psychology: Human Perception and Performance. 1 4: 323–327.

Leibowitz, & H. L. Teuber (Eds.) Handbook of Sensory Physiology: Perception. Berlin: Springer-Verlag.

Malania M., Kezeli A., Kezeli T., Roinishvili M. 2002. “The Double Illusion: A NewType of Color Aftereffect ”, Presented at the Joint Congress, Tubingen, Germany. Published in the Pflügers Archiv — European Journal of Physiology, vol. 443, Supplement 2, 15–19 March, 2002.

Malania, M; Janelidze, D; Roinishvili, M; Kezeli, A. 2004. Induction of McCollough effect causes changes in capasity for constant perception of colour. 27th European Conference on Visual Perception pp. 62, Budapest, 22–26 August, 2004.

McCollough, C. 1965. Color adaptation of edge detectors in the human visual system. Science, 149, 1115–1116.

McCollough, C. 2000. Do McCollough effects provide evidence for global pattern processing? Perception & Psychophysics, 62, 350–362.

Murch, G. M., & Hirsch, J. 1972. The McCollough effect created by complementary afterimages. The American Journal of Psychology, 85, 241–247.

Sharikadze M. & Kezeli A. 2002. Mechanisms of Compensation and Correction of Visual Information are Spatially Organized. Pflügers Archiv, European Journal of Physiology, Vol. 443, Suppl.

Sheth, B. R., & Shimojo, S. 2008. Adapting to an aftereffect. Journal of Vision, 8(3, Art. 29), 1–10.

Siegel, S., & Allan, L. G. 1987. Contingency and the McCollough effect. Perception & Psychophysics, 42, 281–283. Stromeyer, C. F. 1969. Further studies of the McCollough effect. Perception & Psychophysics, 6, 105–110. Stromeyer, C.F. 1978. Color aftereffects dependent on form. In R. Held, H.W.

White, K. D.; Petry, H. M.; Riggs, L. A.; Miller, J. 1978. “Binocular interactions during establishment of McCollough effects”. Vision Research. 18 9: 1201–1215.

Глава 5.
Нейрофизиология зрения и нейротехнологии целенаправленного поведения

Ю. Е. Шелепин, В. Н. Чихман

Давая определение термину «нейротехнологии», большинство исследователей считают, что нейротехнологии — это или совокупность технологий, созданных на основе принципов функционирования нервной системы и органов чувств или технологии исследования и моделирования сенсорных и высших психических функций, необходимых для понимания деятельности мозга. Но оба эти определения не совсем точны и не в полной мере отражают суть происходящей технологической революции. Нейротехнологии представляют собой новый класс стратегических интеллектуальных технологий, обеспечивающих построение самообучающихся систем искусственного интеллекта, способных работать с семантикой, а также реализовывать самостоятельное целенаправленное поведение. В ближайшие годы именно они обеспечат быстрое развитие экономики, социальной и оборонной сфер тех стран, которые планируют и эффективно финансируют ключевые направления науки в области искусственного интеллекта и нейрофизиологии, успешно доводят фундаментальные исследования до создания новых технологий, а новую продукцию — до потребителя. Необходимо подчеркнуть, что наблюдаемый бум нейротехнологий оказался возможным благодаря новым инженерным решениям, реализованным во взаимодействии с исследованиями по физиологии, и в первую очередь, с сенсорной физиологией. Особенно успешным является взаимодействие результатов исследований сенсорных систем с видео, аудио, информационными технологиями, оптикой и, конечно, с оптогенетикой. При этом вклад исследований в области физиологии зрения в успешное развитие различных направлений нейротехнологий является весьма значительным. Особое значение приобретает развитие интеллектуальных технологий распознавания образов и искусственного интеллекта, построенных на основе искусственных нейронных сетей в основе которых лежат общие принципы заложенные еще

в начале двадцатого века в России в Санкт-Петербургской Императорской Военно-медицинской академии и Казанском университете.

Отметим важнейшие моменты в анатомических и физиологических исследованиях, оказавшие влияние на всю современную картину нейронаук и нейротехнологий. Еще в 1890–1893 гг. В. М. Бехтерев, будучи профессором Казанского университета, создал уникальную схему основных проводящих путей головного и спинного мозга человека. Эта двумерная схема была построена на основании огромного количества гистологических срезов головного мозга и последующей мысленной реконструкции проводящих путей. Работу облегчила созданная В. М. Бехтеревым проволочная модель.

Эта проволочная модель по сути является первой в мире 3D моделью архитектуры жестких крупномасштабных нейронных связей в головном мозге человека. Она облегчила В. М. Бехтереву реконструкцию многочисленных гистологических срезов головного мозга человека и до сих пор бережно хранится в Казанском государственном медицинском институте. Двумерный рисунок этой 3D модели (см. рис 1А) вошел в книгу В. М. Бехтерева, изданную в 1893 году [Бехтерев, 1893]. Книга была переведена практически на все основные языки и издана в крупнейших странах мира. Долгие годы книга В. М. Бехтерева была основой клинической диагностики и понимания целостной организации мозга. Спустя сто лет с появлением новых компьютерных технологий оказалось возможным осуществить реконструкцию проводящих путей головного мозга у живого человека, отождествить нейронные связи с основными анатомическими образованиями мозга, с цитоархитектоническими зонами коры (рис.1Б) и даже проследить связь между функционально активными структурами мозга, выявляемыми при выполнении определенных задач, поставленных перед субъектом (испытуемым). Сравнение «Схемы проводящих путей», построенной В. М. Бехтеревым (рис. 1А), и модели (рис. 1Б), являющейся результатом выполнения грандиозной программы «Коннектом», показывает насколько точно В. М. Бехтерев установил основные проводящие пути головного мозга на основании кропотливой работы с тонкими многочисленными прокрашенными срезами извлеченного из черепа и фиксированного в многочисленных растворах головного мозга. Теперь

морфологию проводящих путей головного и спинного мозга можно определять у любого гражданина пришедшего на обследование в крупную клинику. Более того, в специализированных центрах можно пока достаточно грубо оценивать и функциональное состояние и даже функциональные связи в процессе деятельности человека [Шелепин и др., 2009; Шелепин, 2017].

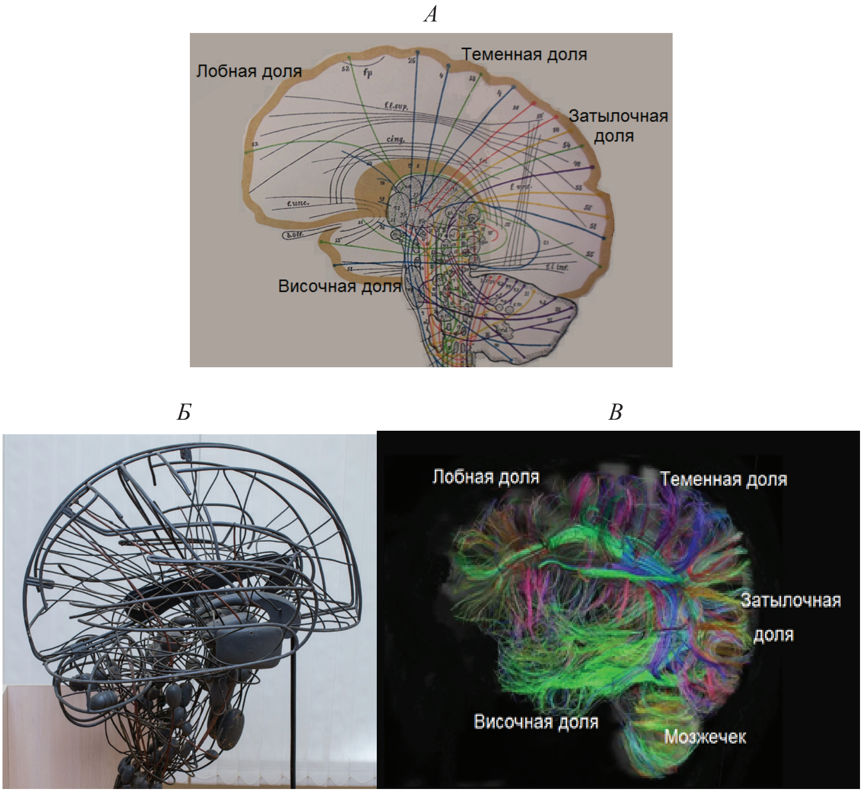
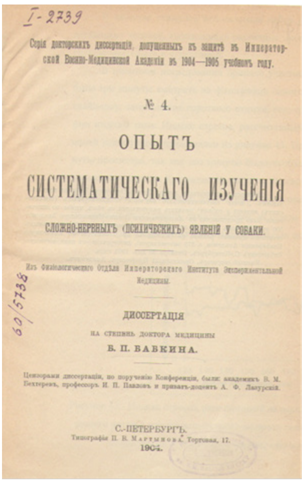


Рис. 1. Реконструкции проводящих путей головного мозга человека: *А* — схема проводящих путей головного мозга, представленная на двумерной плоскости В. М. Бехтеревым, [Бехтерев, 1893]. *Б* — физическая проволочная модель мозга, созданная В. М. Бехтеревым (фото сделано директором Музея истории КГМУ Ивановой Р. Г. при активном содействии профессора А. Л. Зефирова и ассистента А. В. Мартынова (кафедра физиологии КГМУ), и сотрудника КГУ В. Н. Антипова), *В* — реконструкция связей в результате реализации части проекта Коннектом. (https://en.wikipedia.org/wiki/Connectome)

В 1904 г. Б. П. Бабкин разработал концепцию о временной связи в нейронных сетях живых организмов. Знаменитый ученик И. П. Павлова и В. М. Бехтерева в своей диссертационной работе (рис. 2) «Опыт систематического изучения сложно-нервных (психических) явлений у собаки» [Бабкин, 1904] предложил схему установления временной связи вследствие образования «проторенного пути проведения» между корковым центром условного раздражителя и вкусовым, связанным со слюноотделительным центром продолговатого мозга. Это «проторение пути» является основой всей конструкции обучения искусственных нейронных сетей (см. главу Р. О. Малашина в данной книге). Б. П. Бабкин указывал, что при угасании наоборот «происходит разрыв проторенного пути между высшими центрами». В данной работе не только были заменены старые термины, например, «психический» на новые — «сложно-нервный», но установлены фундаментальные факты, касающиеся процессов угасания и восстановления пищевого процесса. Таким образом, Б. П. Бабкин «оживил» модель В. М. Бехтерева, ввел принцип выработки условного рефлекса, т. е. обучения. Интересно, что все годы учебы в Военно-медицинской академии Б. П. Бабкин провел на кафедре В. М. Бехтерева, а в адъюнктуру поступил к И. П. Павлову.


I-2739

Серія докторскихъ диссертацій, допущенныхъ къ защитѣ въ Императорской Военно-Медицинской Академіи въ 1904—1905 учебномъ году.

№ 4.

ОПЫТЪ
СИСТЕМАТИЧЕСКАГО ИЗУЧЕНІЯ
СЛОЖНО-НЕРВНЫХЪ (ПСИХИЧЕСКИХЪ) ЯВЛЕНІЙ У СОБАКИ.

Изъ Физіологическаго Отдѣла Императорскаго Института Экспериментальной Медицины.

60/5738

ДИССЕРТАЦІЯ
НА СТЕПЕНЬ ДОКТОРА МЕДИЦИНЫ
Б. П. БАБКИНА.

Цензорами диссертаціи, по порученію Конференціи, были: академикъ В. М. Бехтеревъ, профессоръ И. П. Павловъ и приватъ-доцентъ А. Ф. Лазурскій.

С.-ПЕТЕРБУРГЪ.
Типографія П. П. Мартынова. Торговая, 17.
1904.

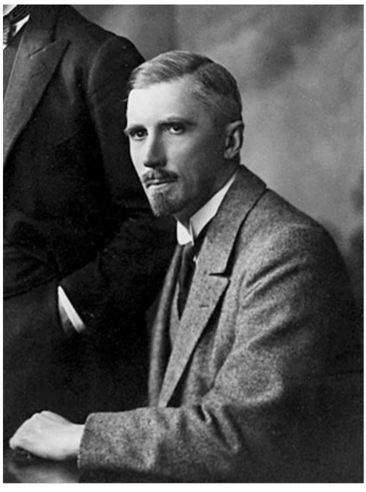


Рис. 2. Слева — титульный лист диссертации Б. П. Бабкина. Справа — портрет Б. П. Бабкина [B. P. Babkin FRS, M.D., D. Sc, LL.D, 1877–1950].

На титульном листе диссертации видно, что цензорами работы выступали совместно В. М. Бехтерев, И. П. Павлов и А. Ф. Лазурский. Труды И. П. Павлова, Б. П. Бабкина и В. М. Бехтерева предопределили основные направления исследование мозга в XX веке. В. М. Бехтерев, И. П. Павлов уверенно занимают свое место в мировой науке. Роль Б. П. Бабкина в развитии мировой науки становится ясной именно теперь в период развития нейротехнологий и моделирования нейронных сетей головного мозга человека. Хотя вихри революции унесли Б. П. Бабкина от родных берегов, он оказался прадедом современной североамериканской научной школы. В начале 20 годов в Великобритании он восстановил утерянные степень доктора наук, профессора, а затем был избран членом Лондонского Королевского общества. Долгие годы вплоть до своей кончины он работал в Канаде, в Монреале, возглавляя кафедру физиологии. Воссоздал там «Башню молчания», вместе с приезжавшим к нему И. П. Павловым и учениками. Перевел на английский язык книгу И. П. Павлова «20 летний опыт изучения высшей нервной деятельности» [Павлов, 1924]. Эти 20 лет Павлов отсчитывал именно от диссертации Бабкина. В Канаде к Б. П. Бабкину пришел учиться учить и «перестраивать мозг человека » двадцатилетний Дональд Хебб, будущий «отец» современной нейрофизиологии и нейротехнологии. Несмотря на свою молодость Д. Хебб был уже директором школы для сложных детей и его интересовали новые методы обучения. Б. П. Бабкин в продолжении несколько десятилетий воспитывал самого Дональда Хебба рекомендовал ему решать то одну то другую задачу, удачно посылал и его и сестру Хебба, тоже ученицу Бабкина стажироваться в разные лаборатории и клиники Северной Америки и Европы. В итоге Дональд Хебб написал известную книгу, в которой изложил основные постулаты деятельности работы нервных сетей. В основе постулатов Хебба лежит временная связь открытая Борис Петровичем в лаборатории Ивана Петровича Павлова и под руководством Ивана Петровича. Современное состояние исследования и создания искусственных нейронных сетей в данной коллективной монографии посвящены специально две главы в которых подробно рассмотрены эти проблемы.

Ввиду технологической направленности данной книги период многолетнего «послепавловского» изучения механизмов высшей нервной деятельности в лабораториях Института физиологии им. И. П. Павлова РАН, мы рассматриваем лишь как узкую часть обширных исследований, проведенных во второй половине двадцатого века в области физиологии зрения. Развитие техники связи, создание вещательного телевидения, радиолокации, эхолокации способствовало интенсивному исследованию сенсорных систем, как систем приема и восприятия информации [Zvorykin 1934–1947; Cherry, 1956]. Изучение организации нейронных сетей зрительной системы позволило связать характеристики зрительной системы со статистическими свойствами реального мира [Глезер, Цукерман, 1961; Глезер, 1985; Цуккерман, 1975; Цуккерман, 1978; Field, 1987, 1993, 1999]. Было показано, что функциональные единицы зрительной коры образуют упорядоченные структуры — модули, пространственно-частотные характеристики которых были измерены. Была определена пространственная организация и функциональные свойства модулей стриарной коры с точки зрения осуществления ими пространственно-частотного описания воспринимаемых изображений. На основе проведенных исследований были разработаны принципы возможного использования описания изображений на уровне стриарной коры нейронными сетями престриарной, нижневисочной и теменной коры, в том числе выяснена роль операторов (фреймов) теменной коры в кодировании зрительной пространственной информации [Глезер, 1985]. Спектральный подход к моделированию механизмов зрительного восприятия [Campbell, Robson, 1968] позволил реализовать новый метод в исследовании нейрофизиологических механизмов зрительного восприятия человека в клинике и эргономике — визоконтрастометрию [Шелепин и др., 1985].

Изучение крупномасштабных нейронных сетей позволило развить представление об архитектуре нейронных сетей зрительной системы. Во многом это стало возможным с развитием нейроиконики — науки об изображениях и нейронных механизмах, обеспечивающих представление наблюдаемых изображений в мозгу живых организмов, их восприятие и распознавание. Нейроиконика включает в себя, как цифровую обработку изображений, например, синтез

зрительных стимулов, так и анализ изображений — карт откликов нейронно-глиальной сети головного мозга в ответ на предъявленные наблюдателю стимулы. Термин архитектура связей пришел из лексикона жестко организованных вычислительных систем. В действительности нейронная сеть является динамической, т.е. живой сетью, у которой много общего с биоценозом. В качестве примера можно привести колебания активности нейронов и глии, которые в результате взаимодействия можно моделировать как колебания численности хищников и жертвы [Шелепин, 1970], при этом возмущения, приходящие на эту сеть из сенсорных входов, меняют и режим их взаимодействия. Реакция нейро-глиальных сетей на сенсорный сигнал — одно из наиболее интересных направлений современных исследований.

Для успешного развития нейротехнологий целенаправленного поведения, важен именно пространственно-временной анализ функционирования нейронной сети, вовлеченной в решение поставленной задачи. На основе физических явлений — оптических волн переключения, пространственных и пространственно-временных диссипативных солитонов в нелинейных оптических системах были предсказаны и экспериментально показаны пути создания новых нейроморфных технологий [Розанов, 2011].

Синхронизация импульсных потоков — один из основных механизмов, лежащих в основе кооперативного функционирования систем целостного мозга. Известно, что степень синхронности импульсных сигналов зависит от двух факторов — от параметров тестового стимула и уровня возбуждения нейронов, вызываемого стимулами. Возникает вопрос — какой же из этих факторов является определяющим в «управлении» синхронностью импульсных реакций нейронов? Параметры зрительных стимулов, «согласованные» со структурой рецептивных полей нейронов зрительной системы, или уровень возбуждения нейронов, создаваемый специфическими или неспецифическими для нейронов возбуждающими воздействиями? Данный вопрос требует дальнейшего изучения.

Для согласования результатов нейрофизиологических исследований в области физиологии зрения с результатами математического моделирования были проведены параллельные исследования на

человеке и вычислительной машине. Они были начаты Н. Н. Красильниковым еще в 50-е годы одновременно с Таннером и Светсом для проверки гипотезы о механизмах согласованной фильтрации в решении задач распознавания [Красильников, 1986, 2001, 2011]. Развитый Н. Н. Красильниковым метод согласованной фильтрации позволил построить количественные модели процесса зрительного восприятия с использованием в качестве обоснования основных психофизических законов. Ключевой элемент этих моделей — согласованный фильтр, который строится с использованием заранее заготовленных шаблонов объекта. Для большего подобия тому, что происходит в мозге человека, согласованный фильтр требует обучения. Поэтому модель согласованной фильтрации дополняется, в настоящее время реализована с помощью нейронных сетей глубокого обучения и позволяет оперировать большим объемом данных. О моделях обучаемых нейронных сетей см главу Малашина в данной книге.

Последним этапом обработки зрительной информации является принятие решения. Затем сигнал запоминается и в ряде случаев происходит осознанный и законченный двигательный ответ. Следует заметить, что двигательный ответ на сигнал может происходить и неосознанно. Данная методика была исходно разработана для обнаружения сигнала известной формы, хранящейся в памяти воспринимающей системы. Однако в реальной жизни при зрительном восприятии сложных сцен с разной освещенностью, поворотом, наклоном, степенью перекрытия и зашумленности эти методы оказываются не всегда эффективными. В связи с этим, особый интерес в нейротехнологии целенаправленного поведения представляют разработки адаптивно обучающихся алгоритмов, воспроизводящих оптимальную реакцию человека при анализе видеоданных. В критических ситуациях, требующих оперативного формирования оптимальных решений (например, в сложных экономических, производственных и, особенно в боевых условиях), такие интеллектуальные нейроморфные системы могут контролировать поведение человека и подсказывать ему особенности возникающей ситуации. При успешном решении задачи они могли бы полностью исключить человека из контура принятия решения там, где скорость реакции человека не-

достаточна или имеется неустойчивость нейрофизиологического отклика, способного привести к невозможности успешного выполнения поставленной задачи. Таким образом, исследование механизма принятие решения является одним из ключевых направлений в развитии нейротехнологий. Выбор, как принятие решения, охватывает все виды осознанной, а в ряде случаев и неосознанной деятельности человека. Особый интерес представляет принятие решений человеком в условиях неопределенности — фактор, достигший максимальных возможностей у человека. Всегда трудно выбрать правильный баланс между скоростью принятия решений и достаточностью имеющейся информации. Нельзя не упомянуть в связи с этим, что первый количественный критерий оценки целенаправленной деятельности был предложен академиком А. А. Харкевичем, как мера оценки целесообразности информации — мера Харкевича [Харкевич, 1973]:

$$x = \log \frac{P_1}{P_0},$$

где P_0 — вероятность достижения цели до получения информации; P_1 — вероятность достижения цели после получения и использования информации.

Получение человеком именно зрительной информации обеспечивает в основном достижение цели. Поэтому нейрофизиологические исследования по восприятию зрительных сигналов «изображений-стимулов», стали ключом к выявлению скрытых для понимания конструкций мозга. Исследования в данной области крайне важны, так как обработка зрительной информации в мозге обеспечивает построение зрительной образной картины мира. Зрительная картина мира — это основа конкретного образного мышления человека [Глезер, 1985]. Именно на основании исследований работы человека в зрительной среде можно понять основные принципы работы мозга, на базе которых строится поведение человека, взаимодействие людей, принципы целенаправленного управления человеком коллективами и технологическими процессами. Зрительная система — это своеобразный интерфейс между оптическим и зрительным пространством в нейронных сетях головного мозга. Именно такое

название имел проект ГКНТ СССР «Рекогнитрон» [Шелепин, 1988]. Проект входил в перспективное направление исследований «Новые информационные технологии (Представление знаний в системах человек и машина)», руководил которым академик Г. С. Поспелов. В качестве критерия оценки интеллекта им было предложено считать способность системы формулировать и достигать поставленную цель. Критерий А. А. Харкевича служил определяющей характеристикой.

В настоящее время математические подходы к изучению механизмов принятия решения все больше захватывают психологов [Барабанщиков, 1990, 2002; Шендяпин и др., 2010]. Математические модели принятия решений были предложены еще в восемнадцатом веке выдающимся петербургским математиком Даниилом Бернулли на примере создания теории игр. Цель работы Д. Бернулли состояла именно в развитии теории принятия решений в условиях неопределенности и получении соответствующих формул оценки. Развитие идей Д. Бернулли привело к тому, что Нобелевская премия в 2002 году в области экономики была присуждена двум американским ученым: психологу Даниэлю Канеману и экономисту Вернону Смиту, показавшим, что люди в экономической сфере действуют менее разумно, чем они это предполагали. Д. Канеман показал как ошибки, свойственные людям при принятии решений, влияют на действия человека в экономике, а В. Смит рассматривал иррациональность поведения человека. Премия была присуждена «за обогащение экономической науки результатами психологических исследований, особенно в отношении оценки человеком ситуации и принятия им решений в условиях неопределенности». В условиях неопределенности человек может действовать не рационально, а иррационально, совершая ошибки не случайные, а вполне определенные, систематические. Людям свойственно повторять не успешные стратегии, а те, в которых неудача произошла, по их мнению, из-за совершенной ошибки или неудачных обстоятельств. Чтобы не совершать повторных ошибок следует помнить, что «не за то отец сына бил, что играл, а за то, что отыгрывался». Иррациональное желание отыграться — основа игорного бизнеса, основа военных просчетов и социальных потрясений. Иррациональные действия не

случайны, они могут быть обусловлены архетипами поведения, зачастую неразумными в современных условиях, но возможно выгодными тысячи лет назад. Подобные проблемы привлекают огромное число исследователей в области экономической психологии. Для развития нейротехнологий среди работающих моделей целенаправленного поведения в условиях неопределенности интересны подходы развитые Саати, в частности принципы иерархической модели Саати [Saati, 2008]. В мозге, вероятно, имеются нейронные сети, осуществляющие принятие решения по множественным критериям и по конфликтующим оценкам, напоминающим модели Саати. При выборе стимулов, имеющих различное значение для наблюдателя, активируются определенные участки во фронтальной коре. В исследованиях, проведенных нами, были показаны наглядные различия активности нейронов между дорсолатеральной префронтальной зоной, медиальной префронтальной зоной и орбитофронтальной зоной [Шелепин, 2009, 2017]. Результаты этих работ позволили предположить существование нескольких центров принятия решений во фронтальной области. Интересно отметить, что между этими областями имеются оппонентные внутрикорковые связи, обеспечивающие поведение. Вероятно, правильнее было бы говорить не об областях, а о нейронных сетях, точнее о нейро-глиальных сетях обеспечивающих целенаправленное поведение и принятие решений. Оппонентная конструкция нейро-глиальных сетей фронтальной области обеспечивает адекватное поведение в повседневных и в экстремальных ситуациях. Решения в эмоциональной сфере могут конфликтовать с решениями в распознавании образов. Мозг человека при принятии решений, работает с применением множественных критериев. Действительно, мы одновременно можем принимать решения разного типа — поворачивать руль автомобиля в соответствии с дорожной обстановкой — и при этом делать мысленный выбор по возникшей на службе или в быту проблеме, да еще при этом разговаривая по мобильному телефону совершенно о другой тематике. Все «одновременно» делаемые процессы относительно медленные, они могут квантоваться более быстрыми нейронными механизмами. Заметим, что максимальная скорость квантования в одном нейроне не может превышать 200 Гц. Более быстрое квантование обеспечивают парал-

лельные и многоуровневые нейронные оппонентные системы, завершающиеся в лобных областях коры головного мозга человека и дающие команды в моторные, исполнительные области мозга. В ряде исследований показано, что в основе работы механизмов принятия решений в мозге человека действительно лежит работа так называемых оппонентных систем отдельных нейронов, «ансамблей», включающих различные мозговые структуры.

Принцип оппонентности присущ всем уровням сенсорных систем. Суть его состоит в том, что возбуждение одних нейронов и даже мозговых структур, приводит к торможению других. Наиболее ярко этот принцип проявляется в работе первичных «входных» каскадов зрительной системы. Было показано, что принятие решений происходит также благодаря оппонентным системам, но расположенным на «выходных каскадах». Эти «выходные каскады» находятся на разных уровнях зрительной системы, вплоть до лобных областей. Судя по всему, «оппонентность» при принятии решений была осознана людьми очень давно и именно поэтому еще в древнем Египте была предложена простая модель этого процесса в виде разновесов, названных «Судом Оссириса». Эта модель оказалась устойчивой и непоколебимой на протяжении нескольких тысяч лет, она является, с одной стороны, интуитивно понятным символом юстиции, а с другой стороны, мы только начинаем осознавать, что действительно мозг человека обладает устройством которое схематически можно проиллюстрировать разновесами. При принятии решения мозг располагает на каждой «чаше разновесов» разную информацию, проявляющуюся в активности оппонентных систем, а именно нейронных сетей, участвующих в принятии решения. Дисбаланс оппонентных систем, по-видимому, объясняет пароксизмы непрерывной деятельности — особого состояния человека, при котором принятие решения он осуществить не может. Термин «пароксизмы непрерывной деятельности» был введен Ф. Д. Горбовым [Горбов, 1963], выдающимся психологом, готовившим первых отечественных космонавтов к космическим полетам. При решении задач, требующих концентрации зрительного внимания, необходимого для надежной обработки больших массивов информации и ответственности при оперативном принятии решений, необходима разработка

новых методов исследования поведения и построение новых моделей деятельности человека в пространстве и во времени. Принцип оппонентности обеспечивает как вход сенсорных систем, так и «выход» — принятие решений.

Принцип параллельности и многоуровневой организации, также присущ входным и выходным нейронным сетям зрительного анализатора. Полученные нами данные позволяют понять, как мозг обеспечивает принятие решений по множественным критериям и конфликтующим оценкам. Результаты работ легли в основу предложенного нами метода оценки состояния оператора и позволяют предсказать пароксизмы деятельности человека при принятии ответственных решений в условиях неопределенности. Важно подчеркнуть, что при изменении стоящей перед человеком цели происходит перестройка активности всей крупномасштабной нейронной сети. Так, в последние годы показана возможность перестройки нейрональных сетей (активности связей) в любой части мозга в процессе решения задачи для достижения поставленной цели [Le Magueresse, Monyer, 2013]. Наглядным примером этому служит перестройка деятельности нервной системы и двигательного аппарата при потере нижних конечностей: человек в этом случае для достижения цели начинает ползти, катиться, идти на руках, чтобы достичь поставленную цель. Недаром специалисты в области искусственного интеллекта говорят, что в создаваемом интеллектуальном устройстве должна быть заложена способность к целеполаганию, формированию цели, а как начальный этап — формирование способности воспринимать семантику не жестко сформулированной инструкции по целеполаганию и стремиться выполнять ее. При патологии мозга, если пациент может осознать инструкцию и следовать ей, невзирая на поражения, игнорируя их и находя обходные пути, цель восстановления человека будет достигнута благодаря разработанным в лаборатории новым нейротехнологиям. Поэтому, именно принцип перестройки нейронных сетей мы положили в основу технологии восстановления больного мозга [Грант РНФ 2014, № 14–15–00918].

Указанные важнейшие нейроанатомические и нейрофизиологические решения легли в основу разработки современных нейротехнологий, например, для создания и обучения искусственных нейрон-

ных сетей. В целом, принцип работы данной технологии известен. Отметим только, что в современных моделях деятельности нейронных сетей ключевым элементом является ассоциативное обучение, суть которого сводится к запоминанию обучающих выборок таким образом, чтобы при предъявлении новых данных, нейронная сеть могла дать ответ, какая из запомненных выборок наиболее близка к вновь поступившей информации. Данная особенность является «отражением» принципа работы человеческой памяти, которая ассоциативна по своей природе. Теперь ассоциативные нейронные сети глубокого обучения стали действительно работающим инструментом спустя 30 лет после пионерских работ Тейво Кохонена [Kohonen, 1977], сделавшего решающий шаг от моделей Д. Хебба и Ф. Розенблатта [Rosenblatt, 1962] к практическим технологическим решениям обучения нейронных сетей. Интересными примерами новых нейротехнологий, созданных на основе полученных данных о работе зрительного анализатора — является программный комплекс управления восстановлением психики человека, внедренный в различные клиники [Грант РНФ 2014, № 14–15–00918], а также ассистивные технологии видеоокулографии, которые основаны на регистрации направления взора человека с помощью специализированных высокоскоростных камер. Ассистивный трекер — это незаменимый помощник в кризисных состояниях, сопровождающихся обездвиженностью человека, так как позволяет человеку управлять с помощью взгляда компьютером, что в свою очередь позволяет мгновенно информировать медицинский персонал, родственников о состоянии, дискомфорте, пожеланиях, общаться с окружающим миром, пользоваться интернетом, большинством компьютерных программ и приложений, синтезировать речь (Нейроиконика–ассистив, shelepink@yandex.ru).

Мы указали очень малую часть существующих возможностей применения нейротехнологий в клинике. В действительности нейротехнологии открывают новые, практически безграничные возможности для обеспечения психического и физического здоровья человека. Таким образом, понимание принципов перестройки нейронных сетей чрезвычайно важно при построении архитектуры нейроморфных процессов, обеспечивающих синтез нового знания, соз-

дания новых технологий принятия решений, управления военными, экономическими и социальными процессами. Аналогично решению клинических задач, нейротехнологии порождают новые возможности инструментов эффективного управления в деятельности общества.

Особый интерес представляют механизмы осознанного и неосознанного восприятия зрительных сцен. Вероятно, существуют каналы зрительной системы, которые обеспечивают эти два типа восприятия зрительных сцен. Показано, что эти каналы могут быть использованы для осознанного и неосознанного воздействия на зрителя вещательного телевидения или пользователя интернета, ЭВМ. Сенсорные каналы восприятия пространственной низкочастотной и временной высокочастотной информации имеют принципиальное отличие от каналов, выделяющих высокие пространственные и низкие временные частоты. Одни каналы передают преимущественно неосознанную, а вторые — преимущественно осознанную информацию. Они имеют разную анатомическую и биохимическую организацию и образуют разные нейронные сети от сетчатки до областей лобной коры, завершающих процесс распознавания, принятия решений и при необходимости — организации двигательных команд.

Функциональное построение взаимодействующих нейронных сетей создает паттерны активности, специфичные для каждой экспериментальной ситуации. Большинство этих сетей представляют систему параллельных каналов, имеют общее начало — глаза человека, и общий выход — «моторные» области коры, планирующие и организующие двигательную реакцию. Каналы начинаются в различных ганглиозных клетках сетчатки и заканчиваются в своих нейронных ансамблях двигательной коры. Каждый из каналов имеет свои особенности, каждый из них выполняет свою функцию. Подобная организация предполагает, что мы можем зарегистрировать слабые моторные ответы на осознанные и неосознанные стимулы. Это, в первую очередь, моторные реакции — микродвижения глаз при фиксации взора на выделенном интересном объекте; это эмоциональные микроколебания диаметра зрачка; это реакции мимических мышц и мышц руки, управляемых зеркальными нейронами, и наконец, это управление позой человека. На основании анализа результатов

наших исследований были выявлены психофизиологические маркеры информационно-психологического воздействия на человека при предъявлении зрительных стимулов — лиц, фигур людей и текстов через систему невербальной, а также вербальной коммуникации [Шелепин, 2017].

Нейрофизиологические механизмы, лежащие в основе построения каналов зрительной системы, — рецептивные поля нейронов зрительной коры, которые, в частности, осуществляют вейвлетную фильтрацию. Следует учитывать, что полосовая фильтрация в двумерном пространственно-частотном спектре естественных изображений оказалась эффективной для решения многих практических задач [Чихман и др., 2011; Chikhman et al., 2011]. Например, фильтрация вертикальных составляющих в двумерном пространственно-частотном спектре изображений облегчает на последующих этапах распознавание фигуры человека, а горизонтальных — лица человека [Барабанщиков, 1990–2012]. В результате работы последних лет в лаборатории физиологии зрения установлено, что диагональная фильтрация изображений облегчает выделение эмоциональных выражений лица человека [Logunova et al., 2014]. Мы обращаем внимание на то, что в процессы, связанные с оценкой эмоционального состояния другого субъекта: по мимике лица при чтении текста, рассматривании картин или натуральных сцен вовлечены все каскады обработки зрительной информации. Указанные структуры участвуют в моделировании психического состояния другого субъекта. Полученные нами фМРТ данные показывают, как и по каким каналам организуется работа нейронных сетей, выделяющих области интереса в изображении натуральной сцены, лица, фигуры человека или толпы, и в тексте [Шелепин, 2017].

Представленные в книге экспериментальные материалы указывают на то, что, во-первых, существуют системы совместно работающих зон мозга. Каждая из систем вносит свой специфический вклад в осуществление психических процессов. Иными словами, каждая из этих систем может вовлекаться в реализацию различных функций в зависимости от стоящей перед испытуемым цели, получаемой инструкции, внутренних и внешних мотивов, экспериментальных условий и физических свойств окружающей среды (в том

числе оптических).

Нами разработан ряд аппаратно-программных средств, хорошо зарекомендовавших себя при решении многих практических задач исследования осознаваемого и неосознаваемого восприятия структуры изображений [Чихман, 2003, 2009, 2012; Шелепин, 2017]. Аппаратно-программные средства объединяют систему синтеза тестов и систему цифровой обработки изображений естественных сцен. Имеется возможность фильтрации сложных изображений (текстов, лиц и сцен), создана система семантической обработки текста; система предъявления изображений с любыми угловыми, спектральными и временными характеристиками [Шелепин, 2017; Chikhman et al., 1998, 2006, 2010] Есть возможность реализации психофизических методов измерения порогов восприятия текста, в которых осуществляется контроль регистрации движений глаз, колебаний диаметра зрачка; электромиография лицевых мышц и термография их скрытого напряжения. Разработана система тестов, предназначенная для того чтобы избирательно активировать каждый из каналов зрительной системы, оценки выходных быстрых откликов мозга по ЭЭГ сигналу или медленных реакций фМРТ методом функциональной магнитно-резонансной томографии. Все методы направлены на объективную оценку когнитивных функций и эмоционального состояния человека. Была изучена роль крупномасштабных нейроно-глиальных сетей, обеспечивающих невербальную коммуникацию между людьми. Разработанные и предложенные нами методы исследования реакций человека, а также методы подготовки и синтеза тестовых сигналов с привлечением методов нейроиконики и нейролингвистики, позволили вскрыть механизмы формирования индивидуальных и архетипических образов.

Было высказано предположение, что каждый архетип имеет свою индивидуальную особенность. Хотя отличия этих особенностей незначительны, но их можно выявить, применив специфический для каждого архетипа набор стимулов. Для этих целей нами в качестве объекта исследования и инструмента изучения архетипа был выбран определенный орнамент, а В. Семеновой в Институте физиологии им. И. П. Павлова РАН создана уникальная коллекция орнаментов народов мира. С. В. Прониным был разработан новый

метод обработки изображений орнаментов, аналогичный обработке изображений лица по особенностям пространственно-частотного спектра. Развитие этих исследований может быть интересно для этнографии. Построение орнамента, его «синтез» или дизайн и восприятие может быть одним из направлений нейроэтнографии, которая наряду с уже сложившимися давно оптогенетикой, генетикой поведения, нейролингвистикой, нейроиконикой, нейроэкономикой и даже нейротеологией и нейросоциологией может оказаться весьма эффективным способом исследования мышления человека.

Возможность детектирования сдвига в эмоциональном состоянии испытуемого при неосознаваемом восприятии эмоционально окрашенных стимулов была показана с помощью метода электромиограммы. При этом реакции испытуемых на неосознаваемые эмоциональные стимулы по мощности ответа были сопоставимы с реакциями на осознаваемое предъявление тех же стимулов.

Предложена и реализована модификация методики исследования взаимосвязи статистик динамических изображений и работы нейронных механизмов на базе мультифокальных вызванных потенциалов. Она позволяет изменять электрическую активность мозга с помощью стимуляции зрительного поля таким образом, чтобы она содержала информацию об определенных паттернах, содержащихся в изображении, но невидимых наблюдателю. Это связано с тем, что сигнал подается по каналу приема неосознаваемой информации. Физиологическая оценка осознанного и неосознанного восприятия в настоящее время на существующей технологической базе контактной регистрации электромиограммы или дистанционной с помощью термографии лицевых мышц представляется достаточно информативным методом выявления эмоциональных реакций.

Особое значение для исследования зрительного восприятия представляет изучение характеристик движения глаз в зависимости от ситуации, контекста, инструкции [Барабанщиков, 1990, 2002, 2012]. Если проанализировать скорость микросаккады в последовательности повторных фиксаций взора, сделанных в ходе зрительного поиска, можно обнаружить важные параметры, один из которых связан с визуальным контентом изображения [Шелепин Е., 2018]. Фиксации по целям вызывают больше микросаккад. Больше микро-

саккад генерируется для тех целей, в которых труднее преодолеть неоднозначность. Другой компонент уже связан не со зрительным процессом, а с процессом принятия решений — финалом зрительного восприятия. Фиксации взора, сопровождающиеся в опыте нажатием кнопки «принятие решения», генерировали больше микросаккад, чем фиксации взора на той же цели, но без нажатия кнопки. Расширение зрачка при тех же повторных фиксациях взора показывает подобную модуляцию. Это позволяет нам сделать вывод о том, что общие механизмы активации являются важнейшей и неотъемлемой частью процессов восприятия, определяющих фиксационные движения глаз. Следует отметить, что движения глаз в процессе чтения представляют собой совершенно необычный тип движений. Мы осуществляем сканирование страницы, и этому процессу, как мы знаем, необходимо учиться. Были изучены условия восприятия текста — характеристики знаков, их контраста, размера, яркости и цвета фона. Особое внимание мы уделяли семантике текста. Удается действительно связать семантику текста и физиологические характеристики контроля чтения. Были разработаны методы автоматического анализа эмоциональных составляющих текста — числа прилагательных и глаголов. Исследовали роль прилагательных, обозначающих цвет, и их влияние на эмоциональное содержание и реакции читателей. Впервые был создан метод, позволивший измерить характеристики движений глаз и их взаимосвязь с пониманием текста. Айно Ламминпия показала, что скорость чтения зависит от размера фовеолы читателя, числа захватываемых знаков, что определяет пропускную способность фовеолы, а понимание текста оказалось обратно пропорционально числу захватов [Ламминпия и др., 2018].

Дальнейшее развитие исследований механизмов зрительного восприятия — это продолжение изучения восприятия динамических изображений сцен. В основе понимания сцен существенное значение приобретает семиотика — наука о знаках и знаковых системах. Под знаком понимается некий объект [или явление], имеющий для наблюдателя информационное значение. Проблемы, стоящие перед семиотикой, рассмотрены как физиологические проблемы изучения конструкции «зрительного» мозга. Восприятие знака — работа «зрительного» мозга, в который включены все нейронные структуры,

участвующие в зрительном восприятии, и согласование с обучением. Происходит параллельное и последовательное взаимодействие нейроно-глиальных и сосудистых сетей мозга, каждая из которых вносит свой вклад в обеспечение целенаправленной деятельности человека. Конструкция «зрительного» мозга такова, что его работа направлена именно на поиск и анализ знаков. Мозг осуществляет описание, как простых знаков — следов, так и сложных знаков — символов и понятий. Одни нейронные сети осуществляют анализ объектов и сцен, отношения между знаками и системами знаков, другие нейронные сети описывают значения знаков и их систем в конкретной ситуации. И конечно, особый интерес представляет эффективность использования получаемой наблюдателем информации, анализ нарушения восприятия, мышления и речи, как нарушения анализа знаков и знаковых систем, отношений между знаками и учета ценности знаков для человека, осуществляющего принятие решений в зависимости от конкретных условий.

Параллельные и последовательные действия сетей зачастую представляют как рекурсивные модели.

Наглядный пример рекурсивных структур, изображений дают два зеркала, поставленные друг против друга (Рис.3). Рекурсивные принципы распространены в информационных технологиях, связанных с синтезом и с обработкой изображений в лингвистике и в когнитивных исследованиях. В большинстве действующих структур многоканальной обработки информации заложены пирамидальные, рекурсивные принципы.

Этот подход равно далек как от эквипотенциальности мозга, так и от узкого локализационизма. Мы уже обсуждали, что структуры мозга, обеспечивающие важнейшие когнитивные процессы, взаимодействуют по принципу оппонентности, т. е. осуществляют решения по множественным критериям и по конфликтующим оценкам. Такое взаимодействие может осуществляться и на глобальном (между областями, отвечающими за активное выполнение задачи и пассивное наблюдение), и на локальном (межмодульном взаимодействии в одной области) уровнях. Паттерн активации мозга зависит от инструкции и иных экспериментальных условий. Имеются области, осуществляющие решения по множественным критериям и по кон-

фликтующим оценкам. Например, решения в эмоциональной сфере могут конфликтовать с решениями в сфере познавательной. Центры принятия решений работают с применением множественных критериев как нейронные оппонентные системы. Такая конструкция фрон-

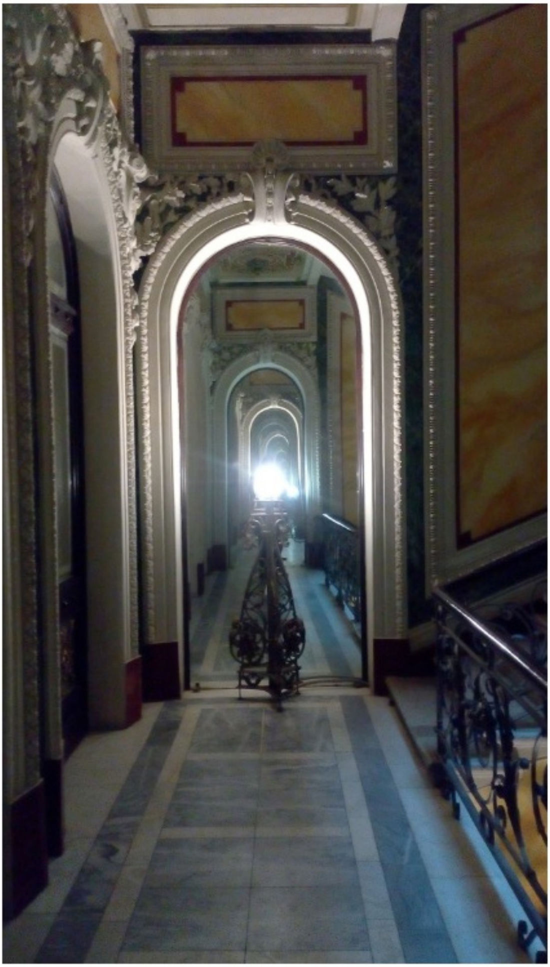


Рис. 3. Рекурсивные изображения зеркал на лестнице Университета ИТМО. (Парадная лестница бывшего особняка купца Елисеева, принадлежащего затем Государственному оптическому институту им. С. И. Вавилова, а теперь Университету ИТМО)

тальной области обеспечивает адекватное поведение в повседневных и в экстремальных ситуациях. Нарушение баланса приводит к «пароксизмам» в деятельности человека [Горбов, 1963; Шелепин, 2017]. Принцип оппонентности является общим. Принцип оппонентности присущ всем уровням сенсорных систем. Суть его состоит в том, что возбуждение одних нейронов и даже мозговых структур, приводит к торможению других. Наиболее ярко этот принцип проявляется в работе первичных «входных» каскадов зрительной системы. Это широкоизвестные, отмеченные Нобелевскими премиями on- и off-системы, цветооппонентные системы. Оппоннетные отношения между ориентационно настроенными нейронами, описывающие форму объекта, менее известны [Шелепин 1981, 1982]. В последние годы эти данные дополнены данными по оппонентному взаимодействию крупномасштабных нейронных сетей [см. главу 9]. В ряде исследований показано, что в основе работы механизмов принятия решений в мозге человека действительно лежит работа так называемых оппонентных систем отдельных нейронов, «ансамблей», включающих различные мозговые структуры [Жукова и др., 2016, 2017; Шелепин и др., 2009–2015; Шелепин и др., 2018]. Оппонентные структуры это будущее развитие искусственных нейронных сетей нового поколения. Новые системы обработки, могут быть построены и частично уже создаются на многоуровневых оппонентных принципах. По мнению Е. Ю. Малаховой и Р. О. Малашина одним из вариантов дальнейшего развития являются нейронные сети типа Adversarial systems, являющимся большим прорывом последних пяти лет в искусственном интеллекте. Adversarial neural networks, работают на принципе соперничества, оппонентности.

В физиологии высшей нервной деятельности Иван Петрович Павлов широко применял представления о возбуждении и торможении, как основы высшей нервной деятельности. Термины возбуждение и торможение в Павловской терминологии были общими как и оппонентные отношения. Интересно их соотношение с принципом дополнительности и с принципом неопределенности в физике 20 века.

Принцип дополнительности — один из важнейших методологических и эвристических принципов науки, а также один из важнейших принципов квантовой механики, сформулированный в 1927 году

Нильсом Бором. Согласно этому принципу, для полного описания квантовомеханических явлений необходимо применять два взаимоисключающих («дополнительных») набора классических понятий, совокупность которых даёт исчерпывающую информацию об этих явлениях как о целостных. Например, дополнительными в квантовой механике являются пространственно-временная и энергетически-импульсная картины. Принцип дополнительности Бор распространил на все явления и биологические и социальные. Вероятно, именно в этом кроется явный интерес Нильса Бора к роли возбуждения и торможения в физиологии высшей нервной деятельности и его дружеские отношения с Иваном Петровичем Павловым. Визит Бора с женой в Колтуши к И. П. Павлову состоялся в начале 30-х годов на пике интереса в обществе и к работам Павлова, и к работам Бора.

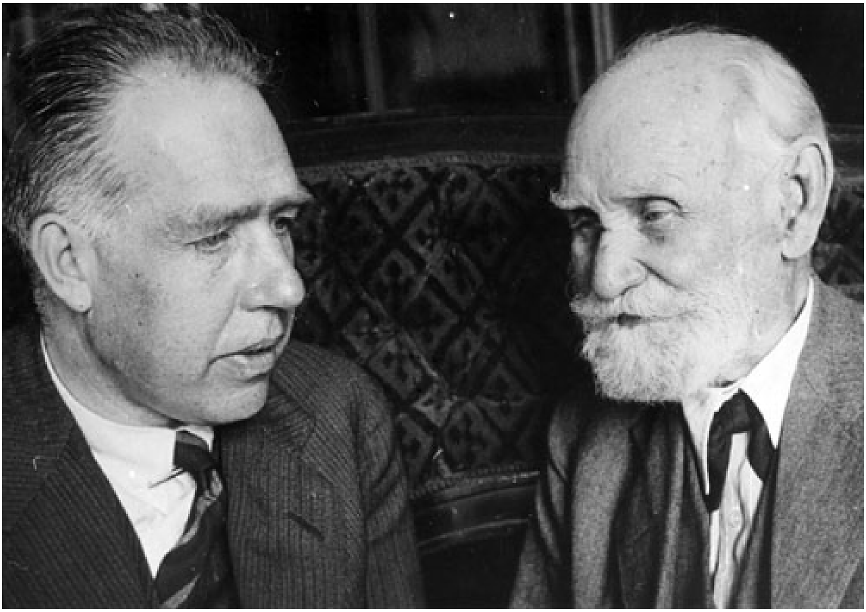


Рис. 4. Павлов и Бор.
(Фото из Архива Института физиологии им. И. П. Павлова РАН)

Приблизительно в эти же годы был открыт принцип неопределённости, Вернером Гейзенбергом (1927 г., основа квантовой меха-

ники), который является следствием принципа корпускулярно-волнового дуализма. Принцип неопределённости, это соотношение неопределённостей, точности одновременного определения пары характеристик исследуемой системы. С тех пор неизменно идет поиск применимости статистических закономерностей, развитых в квантовой механике, к биологии и социологии. Впервые это вопрос был рассмотрен в знаменитой «Книге трех авторов», первый и ведущий из них был Николай Владимирович Тимофеев-Ресовский. Эта книга была написана и издана впервые в Германии в тридцатые годы на немецком языке и лишь недавно переведена на английский. В ней направление и методы для открытия конструкции ДНК были рассмотрены и воплощены в Англии спустя 15 лет. К ней обращаются многие, и она еще в сороковые годы была популяризирована знаменитыми физиками. Этот вопрос в настоящее время связан с важнейшим этапом развития понимания организации мозга. Важно понимание того, как происходит выбор в точке бифуркации в условиях целенаправленного поведения. Загадочный визит Н. Бора с супругой к И.П. Павлову, может быть объяснен совершенно просто. Обоих не устраивал жесткий детерминизм поведения, диктуемый представлениями того времени. Вероятно они обсуждали представления о пластичности нейронных сетей процессы возбуждения и торможения в широком смысле этого слова, представления о роли временной связи между нейронами, введенные в обиход в диссертации 1904 года Б.П. Бабкиным, и развитые его учеником Дональдом Хэббом в постулатах, названных уже его именем.

Принцип оппонентности давно применен был и в лингвистике. Язык отражает и влияет на построение картины мира его носителями. Структура языка определяет структуру мышления и важнейший способ познания внешнего мира. Структурализм в лингвистике возник одновременно с гештальт психологией. Н.С. Трубецкой ввел представление о «бинарных оппозициях» в лингвистике [Трубецкой, 1937, 1960].

Особый интерес представляет изучение текста как гештальта в пространственной, временной и смысловой форме. Такой подход позволяет сравнивать механизмы восприятия текста с механизмами построения гештальта при восприятии изображений сцен, объ-

ектов абстрактных фигур, и даже белого или «окрашенного» шума. На основе бинарных оппозиций созданы методы оценки визуальной информации текстов. Особый интерес представляют методы предсказания в автоматическом режиме их эмоциональных воздействий даже до проведения исследований с большими группами испытуемых. Тексты могут быть проанализированы и синтезированы за счет управления или подбора насыщенности текста теми или иными частями речи (существительные, прилагательные, глаголы).

Есть две области исследований, конвергенция которых, была начата в древности. Работы, проведенные предками на высочайшем интуитивном уровне, дали человечеству разные формы письменного языка. Теперь есть науки — нейроиконика и нейролингвистика. Нейроиконика это наука об изображениях, их построении и восприятии в «зрительном» мозге человека и управлении поведением человека путем синтеза определенных осознаваемых и неосознаваемых изображений, и нейролингвистика как наука о нейрофизиологических механизмах обработки вербальной информации и методах коммуникации и управления поведением человека на основе вербальной информации. Есть различные области совместных интересов этих наук: нейрофизиология исследования невербальных коммуникаций и нейрофизиология механизмов чтения осмысленных и бессмысленных текстов. С одной стороны, тексты являются изображениями, но с другой — эти изображения несут вербальную речевую информацию, являются продуктом и даже частью мыслительной деятельности человека. Тексты — это вполне определенный класс изображений, исследования которых немыслимы без привлечения не только нейроиконики, но и нейролингвистики: ведь для умеющих читать текст это письменная речь. А письменная речь — это удивительная область взаимовлияния изображения и речи. Эти науки неотъемлемая и развитая область нейротехнологий. Чтение представляет собой выработанный в процессе обучения с учителем навык, который обеспечивают разные элементы зрительной и глазодвигательной систем. Этот навык не дается от рождения. В обеспечении процесса чтения участвуют разные этапы обработки зрительной информации и движений глаз. Движения глаз во время чтения — это чередование саккад, фиксаций взора и в этот момент микросаккад. Саккады пе-

ремещают взор по тексту, во время фиксаций происходит распознавание символов, составляющих морфемы и слова. Воспринимаемые символы преобразуются в орфографические и представленные внутренней речью фонологические процессы, которые затем обеспечивают процессы обработки текстовой (языковой) информации, приводящей к пониманию смысла текста.

Движения глаз, зачастую неосознанные движения губ и мимики во время чтения, перераспределение активности мозга — все это физиологические показатели, связанные с реализацией языковых функций. Они представляют собой отражение работы разных механизмов, обеспечивающих чтение и понимание прочитанного. Эти показатели широко используются в психолингвистике как науке о взаимосвязи языка, мышления и сознания.

Анализ текста включает физический, лингвистический, статистический, информационный, психологический и физиологический уровни. Интерес представляет взаимосвязь лингвистических характеристик текста и его психоэмоциональная оценка. Часто для этого исследуют роль прилагательных и глаголов в тексте, а также их роль в определении читателем эмоционального значения текста. «Выход» сенсорных систем — многоканальный и оппонентный, аналогично входу, обеспечивает непрерывную целенаправленную деятельность человека в реальных условиях. Причем он включает как осознаваемые этапы обработки информации, так и неосознаваемые. Оппонентный принцип организации мозга является основой дуалистической структуры семиотики как науки, изучающей взаимодействие всегда двух гностических систем: приемной и передающей, эмоциональной и рациональной, абстрактной и конкретной. Поэтому интерес представляют вопросы — где абстрактное и конкретное описание локализовано и какими нейронными структурами обеспечено. Для развития нейротехнологий необходимо изучение двух важнейших и оппонентных типов описания изображений локальных и глобальных [Шелепин 2008; Chikhman 2012].

Новые нейротехнологии это проблемы конвергенции человека и машины, решение задач автономного управления, целенаправленной деятельности человека и автомата, восстановление утраченных возможностей человека, и создание новых информационных интел-

лектуальных параллельных систем обработки данных, параллельных нейронных систем с разной пространственно-временной полосой пропускания и системой оппонентных механизмов принятия решений, при необходимости подсказывающие человеку оптимальное решение в сложной обстановке.

Литература

Бабкин Б. П. Опыт систематического изучения сложно-рефлекторных [психических] явлений у собаки. [Диссертация, защищенная в Императорской Военно-медицинской академии], СПб, 1904.

Барабанщиков В. А. Динамика зрительного восприятия. М.: Наука, 1990. 240 с.

Барабанщиков В. А. Восприятие и событие. М.: Алетейя, 2002. 512 с.

Барабанщиков В. А. Экспрессии лица и их восприятие. М.: Институт психологии РАН, 2012. 352 с.

Бехтерев В. М. Проводящие пути мозга. Казань, 1893.

Глезер В. Д., Цукерман И. И. Информация и зрение. М-Л, АН СССР, 1961.

Глезер В. Д. Зрение и мышление. Л. Наука. 1985

Горбов Ф. Д. Пароксизмы при непрерывной деятельности: автореф. дис. докт. мед. наук. М.: ИН и ПФ, 1963.

Красильников Н. Н. Теория передачи и восприятия изображений. М.: Радио и связь, 1986. 247 с.

Красильников Н. Н. Цифровая обработка изображений. М.: Вузовская книга, 2001. 320 с.

Красильников Н. Н. Цифровая обработка 2D- и 3D-изображений. СПб.: БХВ-Петербург, 2011.

Ламминпия А. М., Пронин С. В., Шелепин Ю. Е. Пространственно-частотная фильтрация текста для локального и глобаль-

ного анализа. Оптический журнал. 2018. Т 85, № 8 С. 39–45.

Павлов И. П. Двадцатилетний опыт объективного изучения высшей нервной деятельности. Л. 1924.

Розанов Н. Н. Диссипативные оптические солитоны. От микро- к нано- и атто …Изд-во: Физматлит, 2011.

Розенблатт, Ф. Принципы нейродинамики: Перцептроны и теория механизмов мозга — М.: Мир, 1965. — 480с.

Трубецкой Н. С. Цит. по: Основы фонологии / М.: Изд-во иностранной литературы, 1960 [Trubetzkoy N. Essai d'une théorie des oppositions phonologiques. Journal de Psychologie. 1936 Vol. 33. P. 5–18; Trubetzkoy N. Grundzuege der Phonologie. In: Travaux du Cercle Linguistique de Prague. 1939].

Харкевич А. А. Избранные труды в 3-х томах. Том 3. Теория информации. Опознание образов. — М.: Наука, 1973. — 524 с.

Цуккерман И. И. Статистическая структура изображений и особенности зрительного восприятия // Переработка информации в зрительной системе. Л.: Наука, 1975.

Цуккерман И. И. О согласованности пространственно-частотных фильтров зрительного анализатора со статистикой изображений // Биофизика. 1978. Т. 23. № 6. С. 1108–1109.

Чихман В. Н., Шелепин Ю. Е., Солнушкин С. Д., Пронин С. В., Фореман Н., Молодцов В. О., Пржибыш Н. Э., Ляхов О. Н. Методические средства для исследования зрительного восприятия фрагментированных изображений. Российский физиологический журнал, 2003, т. 89, N10, с. 1313–1321.

Чихман В. Н., Бондарко В. М., Голузина А. Г., Данилова М. В., Солнушкин С. Д. Влияние маскировки на опознание контурных изображений. Сенсорные системы. 2009. т. 23. № 1. с. 51–60.

Чихман В. Н., Шелепин Ю. Е., Фореман Н., Пэсмор П. Восприятие фрагментированных изображений трехмерных

объектов при изменении угла наблюдения. Российский физиологический журнал. 2009. т. 95. N4. с. 324–334.

Чихман В. Н., Шелепин Ю. Е., Пронин С. В. Экспериментальное исследование инвариантного восприятия вейвлетных изображений. 2011. Оптический журнал. Т. 78. № 12. С. 50–56.

Чихман В. Н., Бондарко В. М., Голузина А. Г., Данилова М. В., Солнушкин С. Д. Влияние окружения на ошибки опознания при краудинг-эффекте. Сенсорные системы, 2012, т. 26, № 3, с. 195–203

Шелепин Е. Ю. Оптимизация контента сайта в обучающем процессе // Оптический журнал, Том 85, № 8 стр. 95–98, 2018 г.

Шелепин К. Ю., Труфанов Г. Е., Фокин В. А.; Васильев П. П.; Соколов А. В. Цифровая визуализация активности нейронных сетей головного мозга человека до, во время и после инсайта при распознавании изображений // Оптический журнал. Том 85, № 8, стр. 29–38, 2018.

Шелепин Ю. Е. К математической интерпретации нейроно-глиальных взаимоотношений // Доклады Академии наук СССР. 1970. Т. 192. № 3. С. 698–701.

Шелепин Ю. Е. Ориентационная избирательность и пространственно-частотные характеристики рецептивных полей нейронов затылочной коры кошки // Нейрофизиология. 1981 а. Т. 13. № 3. С. 227–232.

Шелепин Ю. Е. Фильтрационные свойства рецептивных полей нейронов зрительной коры // Доклады Академии наук СССР. 1981 б. Т. 261. № 6. С. 1506–1509.

Шелепин Ю. Е., Колесникова Л. Н., Левкович Ю. И. Визоконтрастометрия. Наука. 1985

Шелепин Ю. Е. 1988. «Рекогнитрон» проект № 157, направление «Новые информационные технологии [Представление знаний в системах человек и машина]». Руководитель Г. С. Поспелов. ГКНТ СССР, 1988–1992гг

Шелепин Ю. Е., Чихман В. Н.. Инвариантность зрительного восприятия Экспериментальная психология. 2008. № 1 с. 7–33

Шелепин Ю. Е., Чихман В. Н., Финеман Н. Анализ исследований восприятия фрагментированных изображений — целостное восприятие и восприятие по информативным признакам // Российский физиологический журнал. 2008. Т. 94. Вып.7. С. 758–776.

Шелепин Ю. Е., Фокин В. А., Хараузов А. К., Пронин С. В., Чихман В. Н. Локализация центра принятия решений при восприятии формы зрительных стимулов. // Доклады Академии Наук, 2009, том 429, № 6, с. 1–3.

Шелепин Ю. Е., Чихман В. Н., Пронин С. В., Фореман Н. Локальный и глобальный анализ в зрительной системе. Глава в монографии «Современная психофизика». Изд-во Институт психологии РАН. 2009. Москва. 542 с. ISBN978–5–9270–0151–4.

Шелепин Ю. Е. Введение в нейроиконику, СПб, 2017

Шендяпин В. М., Барабанщиков В. А., Скотникова И. Г. Уверенность в решении: моделирование и экспериментальная проверка // Экспериментальная психология, 2010, том 3, № 1, с. 30–57

Campbell F. W., Robson J. G. Application of Fourier Analyses to the Visibility of Gratings // Journal of Physiology (London). 1968. Vol. 197. P. 551–566.

Cherry C. On human communication. N.Y. 1956

Chikhman V., Shelepin Y., Pronin S., Harausov A., Krasilnikov N., Makulov V. Electrophysiological measurements of the natural image distortion // SPIE. 1998. Vol. 3299. P. 510–518.

Chikhman V., Shelepin Y., Foreman N., Merkuljev A., Pronin S. Incomplete figure perception and invisible masking // Perception. 2006. Vol. 35. No. 11. P. 1441–1457.

Chikhman V. N., Shelepin Y. E., Foreman N., Passmore P. Perception of fragmented Images of Tree-Dimentional Objects as the Ob-

servation Angle Changes. Neuroscience and behavioral Physiology. 2010. v. 40, N5, p. 565–572

Chikhman V., Shelepin Y., Pronin S. Experimental study of invariant perception of wavelet images // Journal of Optical Technology. 2011. Vol. 78 (12). P. 803–807.

Chikhman V.N., Bondarko V.M., Danilova M.V., Goluzina A.G., Shelepin Y.E. Complexity of images: Experimental and computational estimates compared // Perception. 2012. Vol. 41. P. 631–647

Gabor D. Lectures on communication theory. MIT, Cambridge, 1951

Field D.J. Match filters, wavelets and the statistics of natural scenes // Journal of Optical Technology. 1999. Vol. 66 (9). P. 788–796.

Field D.J. Relations between the statistics of natural images and the response properties of cortical cells // Journal of the Optical Society of America. 1987. Vol. 4. P. 2379–2394.

Field D.J. Scale-invariance and self-similar ‘wavelet’ transforms: an analysis of natural scenes and mammalian visual systems. Wavelets, Fractals and Fourier Transforms: New Developments and New Applications. 1993. P. 151–193.

Kohonen T.K. Associative Memory—A System-Theoretical Approach, Springer Series 1977.

Le Magueresse C. and Monyer H., GABAergic Interneurons Shape the Functional Maturation of the Cortex, Neuron 77, February 6, 2013, p 388–405.

Logunova E.V., Pronin S.V., Shelepin Yu. E. Modelling the operation of spatial-frequency filters during the perception of complex dynamic scenes // Journal of Optical Technology. 2014. 81 [11], P. 665–670.

Rosenblatt F., Principles of Neurodynamics. Spartan books Washington, 1962,

Saaty, T.L. Decision making with the analytic hierarchy process // Int. J. Services Sciences, Vol. 1, No. 1, pp.83–98, 2008

Zvorykin V. K. The iconoscope — A modern version of the electric eye // Proceedings of the IRE. 1934. Vol. 22. P. 16–32.

Zvorykin V. K. Iconoscopes and kinescopes in television // RCA Review. 1936. Vol. 1.P. 60–84.

Zvorykin V. K. Automatic recording spectroradiometer for cathodoluminescent materials

// Journal of the Optical Society of America. 1939 a. Vol. 29. February. P. 84–91.

Zvorykin V. K. The electrostatic electron multiplier // Proceedings of the IRE. 1939 b.Vol. 27. September. P. 558–566.

Zvorykin V. K., Flory L. R. An Electronic Reading Aid to the Blind. 1946 // Proceedings of the American Philosophical Society. 1947. Vol. 91. No. 2.

Глава 6.
Пространственные элементы зрительной системы. Моделирование зрительного восприятия

В. М. Бондарко

Описание зрительных объектов и их свойств осуществляется с помощью функциональных элементов зрительной системы, которые организованы в высокоупорядоченные нейронные сети сетчатки, подкорковых ядер и зрительной коры. Как правило, модели зрительного восприятия используют различного рода функциональные элементы для обработки только определенного класса изображений. Поэтому наблюдается их большое разнообразие. Разнообразие моделей связано еще и с тем, что в них рассматриваются разные взаимодействия между элементами. Перестройка нейронных сетей может быть обусловлена как видоизменением самих элементов, так и связей между ними. Наиболее сильные изменения происходят в онтогенезе, а также наблюдаются в процессе обучения конкретной зрительной задаче. В настоящем разделе будут кратко изложены некоторые аспекты этой проблематики.

В нейрофизиологических исследованиях, выполненных на животных, довольно подробно изучены характеристики и предполагаемые функции отдельных элементов или рецептивных полей нейронов более низких уровней зрительной системы — сетчатки, наружного коленчатого тела и стриарной коры. Однако при изучении механизмов зрительного восприятия человеком часто возникает необходимость сопоставлять данные нейрофизиологических и поведенческих экспериментов, полученные на животных, с результатами психофизических или психофизиологических исследований. Для этого во многих случаях необходимо знать конкретные свойства рецептивных полей нейронов зрительной системы человека. С этой целью вводятся в рассмотрение так называемые пространственные элементы. По своему содержанию понятие пространственного элемента близко к понятию рецептивного поля с той лишь разницей, что пространственные элементы отражают реакцию всей зрительной системы на какой-либо стимул. Функционирование каждого

пространственного элемента можно рассматривать как суммарную реакцию рецептивных полей нейронов различных уровней, имеющих близкие свойства и обрабатывающих один и тот же участок поля зрения. Совокупность пространственных элементов, располагающихся в разных участках поля зрения, описывает зрительные изображения. Это описание затем анализируется более высокими уровнями зрительной системы и используется в моделях.

Остановимся сначала на работах, выполненных по изучению пространственных элементов, а затем перейдем к исследованиям возможной перестройки некоторых элементов, происходящей в онтогенезе или при обучении зрительной задаче.

Большое число психофизических исследований посвящено выявлению пространственных элементов, лежащих в основе восприятия формы объектов. Наиболее интенсивно это направление исследований стало развиваться после основополагающих работ Хьюбела и Визела [Hubel, Wiesel, 1962] и Кэмпбелла, Робсона, Блэйкмора [Campbell, Robson, 1968; Blakemore, Campbell, 1969]. В первых нейрофизиологических работах Хьюбела и Визела были изучены свойства рецептивных полей стриарной коры и их организация. В частности, в них были описаны рецептивные поля — «детекторы» полос и краев. Это послужило толчком для нахождения элементов — «детекторов» полос и краев — в психофизике. Работы Кэмпбелла, Робсона и Блэйкмора изначально были выполнены психофизическими методами. На основании полученных данных авторы высказали гипотезу о пространственно-частотном анализе изображений, осуществляемом зрительной системой с помощью набора пространственно-частотных каналов. Кэмпбелл и его коллеги [Campbell et al., 1969] предприняли первую попытку для нахождения нейрофизиологических коррелятов этих каналов. По их мнению и мнению других исследователей такими коррелятами являются рецептивные поля стриарной коры [Глезер, 1985, 1993; Шелепин и др., 1992]. Вследствие этих исследований появилась возможность корректного моделирования процесса зрительного восприятия и сопоставления нейрофизиологических и психофизиологических данных.

Первые работы по такому сопоставлению были выполнены Куликовским, Кинг-Смитом, Шепли и Толхурстом [Kulikowski, King-

Smith, 1973; Shapley, Tolhurst, 1973]. Ими были получены и описаны «детекторы» или пространственные элементы краев, полос и решеток. В своих исследованиях [Бондарко и др., 1983; Бондарко, Гаузельман, 1987, 1988; Бондарко и др., 1999а, б] мы также рассмотрели ряд вопросов, касающихся организации и функционирования пространственных элементов. Нами было показано, что в зависимости от задачи, стоящей перед наблюдателем, эти элементы могут быть различными и с разными свойствами. Так, при обнаружении высокочастотных решеток выявляются пространственные элементы — детекторы полос (рис.1, а), настроенные на оптимальные средние частоты зрительной системы человека. Отметим, что подобные пространственные элементы были описаны ранее Вилсоном и Гелбом [Wilson, Gelb, 1984] при выполнении другой зрительной задачи. Эти пространственные элементы имеют весовые функции, представимые в виде разности двух или трех гауссиан, и по своим свойствам близки к простым рецептивным полям стриарной коры — «детекторам» полос. Нами были выявлены и другие пространственные элементы при опознании тех же стимулов. Ими оказались высокочастотные детекторы решеток с весовыми функциями — элементами Габора с большим числом перемен знака (рис.1, б). По своим свойствам такие элементы подобны сложным рецептивным полям стриарной коры. Отметим, что именно стриарная кора имеет рецептивные поля с высокой пространственной частотой настройки в отличие от других корковых областей, что дает нам право сделать подобный вывод.

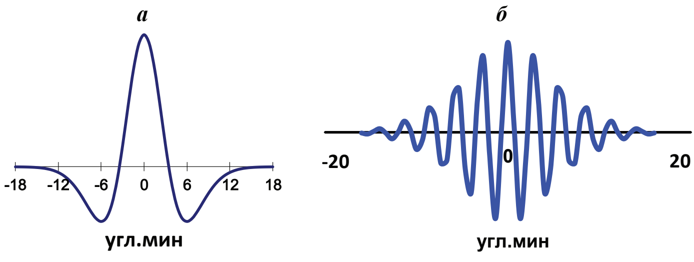


Рис. 1. Весовые функции детектора узкой полосы (а) и детектора решетки (б)

Важным фактом, полученным нами, является то, что зона действия пространственных элементов, служащих для опознания реше-

ток равна 0.4–0.8 угл.град., что соответствует размеру центральной части фовеолы (около 0.5 угл.град.). Примерно такой же размер имеет и «детектор» полосы — он также охватывает центральную часть фовеолы. Такая согласованность в размерах позволяет данным элементам взаимодействовать при решении более сложных задач, связанных с опознанием зрительных изображений. Она может являться основой для использования при описании преобразований типа разложений их в обобщенные ряды Фурье, поскольку пространственные элементы при одинаковом размере настроены на разные пространственные частоты. Эта согласованность будет рассмотрена нами позднее при анализе моделей зрительного восприятия. Особенно хотелось бы отметить также полученную нами экспериментально дискретность настройки пространственных элементов по частотам. С помощью подобных пространственных элементов может осуществляться фильтрация любых изображений. В задачах различения и обнаружения простых стимулов механизмом взаимодействия между элементами может выступать механизм вероятностной суммации, учитывающий независимое функционирование каждого элемента.

Рассмотрим примеры функционирования отдельных элементов и взаимодействия между этими элементами, возникающие при решении более сложных зрительных задач. Одна из таких задач — краудинг-эффект — ухудшение восприятия, возникающее при близком расстоянии между изображениями. Несмотря на относительно длительный период исследования краудинг-эффекта, до сих пор не существует единого мнения о механизмах, лежащих в основе этого феномена [Atkinson, 1991; Flom, 1991; Levi, 2008]. Возможно, это связано с тем, что в зависимости от используемых изображений и условий их предъявления могут функционировать разные механизмы зрительной системы. Так, при изучении краудинг-эффекта на пределе разрешения зрительной системы в наших работах [Бондарко, Данилова, 2002; Данилова, Бондарко, 2002; Danilova, Bondarko, 2007] было показано, что ухудшение восприятия простирается на различные расстояния между тестовым стимулом и окружением в зависимости от используемых стимулов и окружения. Было высказано предположение, что это расхождение может быть связано с преимущественным использованием для описания изображений раз-

личных рецептивных полей нейронов первичной зрительной коры, наилучшим образом осуществляющих задачу различения тестовых объектов с окружением. При изучении краудинг-эффекта у взрослых наблюдателей, нами было показано, что краудинг-эффект с использованием тестовых колец Ландольта может быть объяснен функционированием самых высокочастотных пространственных элементов, образующих высокочастотный канал. Этими пространственными элементами являются детекторы полос и краев. В случае тестовых решеток при окружении их решетками со случайно меняющейся ориентацией такими пространственными элементами, скорее всего, являются высокочастотные детекторы решеток, нечувствительные к изменению фазы решетки. Коррелятами этих пространственных элементов могут служить высокочастотные сложные рецептивные поля. Весовые функции всех высокочастотных элементов согласованы с оптикой глаза и расположением фоторецепторов на сетчатке [Шелепин, Бондарко, 2002].

Исследование краудинг-эффекта с использованием тестовых колец Ландольта было продолжено нами в онтогенезе у детей в возрасте от 3 до 12 лет, у подростков и у взрослых [Семенов и др., 2000; Бондарко, Семенов, 2005]. Было показано, что расстояния между тестовыми стимулами и окружающими изображениями, при которых наблюдается ухудшение восприятия, значительно больше у маленьких детей. Затем эти расстояния уменьшаются с увеличением возраста, и около 9 лет становятся таким же, как и у взрослых. Из морфологии известно, что формирование стриарной коры заканчивается к семи-девяти-летнему возрасту. Таким образом, наши данные подтверждают предположение о локализации механизма краудинг-эффекта на пределе разрешения зрительной системы именно в стриарной коре. Напомним, что ранее Фломом и др. [Flom et al., 1963b] при предъявлении стимула одному глазу, а окружения другому было показано, что механизм краудинг-эффекта локализован на уровне не ниже стриарной коры, поскольку наблюдалось ухудшение на том же расстоянии, что и при предъявлении этих изображений одному глазу. По полученным данным (измерена острота зрения и краудинг-эффект) были рассчитаны весовые функции высокочастотных элементов — детекторов полосы у наблюдателей различного возраста. На

рис. 2 показаны для сравнения весовые функции пространственных элементов у детей 3–6 лет и у взрослых. Весовые функции у детей отличаются от весовых функций взрослых тем, что у взрослых сужаются возбудительная и тормозные зоны, кроме того появляются дополнительные боковые растормаживающие зоны. Весовые функции высокочастотных пространственных элементов у детей представимы в виде разности двух гауссиан, а весовые функции взрослых — трех гауссиан. На языке пространственно-частотного анализа полученное изменение весовых функций означает сужение частотной настройки самого высокочастотного канала у взрослых по сравнению с детьми и настройку пространственных элементов на более высокую частоту.

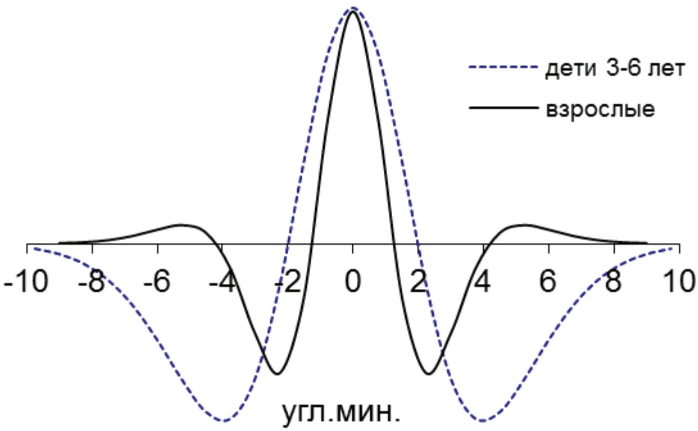


Рис. 2. Весовые функции высокочастотных пространственных элементов — «детекторов» полос у детей 3–6 лет и у взрослых.

Исследование краудинг-эффекта было продолжено с использованием разных изображений, где тестовые кольца Ландольта были окружены полосами, а решетки и буквы Е — решетками и буквами Е соответственно [Бондарко, Семенов, 2005, 2012]. Было показано ухудшение восприятия в случае расположения окружающих изображений вблизи тестовых объектов. Однако возрастные зависимости были разными для колец Ландольта, букв Е и решеток. Расстояния между стимулом и окружением, при которых происходило ухудше-

ние, уменьшались с увеличением возраста для решеток до 16–17 лет, а для других стимулов до 12 лет. Такие различные зависимости можно интерпретировать, как с точки зрения возрастных изменений в соответствующих пространственных элементах, так и их взаимодействия. Известно, что между пространственно-частотными каналами наблюдаются тормозные и растормаживающие взаимодействия [Polat, Sagi, 1993]. Характер взаимодействий зависит от расстояния между изображениями и их пространственно-частотной настройки.

Таким образом, в результате наших исследований в онтогенезе мы получили свидетельства об изменении самых высокочастотных элементов зрительной системы. Настройка пространственных элементов становилась более высокочастотной и узкополосной. Отметим, что сходное поведение наблюдалось и при обучении зрительной задаче — обнаружении узкой полосы. После тренировки пространственные элементы стали более чувствительными и узкополосными [Бондарко и др., 1999a].

Остановимся теперь на вопросе о взаимодействиях между пространственными элементами или рецептивными полями.

В 1978 г. В. Д. Глезером [Глезер, 1978] была высказана гипотеза о том, что рецептивные поля нейронов стриарной коры могут осуществлять так называемое кусочное преобразование Фурье. Под кусочным преобразованием Фурье подразумевалось разложение изображений в конечный набор коэффициентов ряда Фурье, в котором используется не более четырех гармоник (синусов и косинусов). Нейрофизиологическое обоснование возможности такого разложения было получено в работах ряда исследователей (см. например, обзоры в монографиях В. Д. Глезера, 1985, 1993, частично эти работы приведены ниже). Набор рецептивных полей, обрабатывающих один и тот же участок поля зрения и настроенных на различные частоты, был назван модулем [Глезер, 1985]. Были сделаны попытки аппроксимировать результаты психофизических экспериментов моделью модулей.

В первых работах [Вол, Павловская, 1986; Vol et al., 1990] в модели был использован полный Фурье-спектр изображений, а не коэффициенты конечного ряда Фурье. Измерялось Евклидово расстояние между спектрами изображений. Вычисленные расстояния срав-

нивались с так называемыми матрицами смешиваемости изображений или матрицами ошибок (матрица количества ответов, когда i-е изображение называют j-м изображением), регистрируемых в условиях дефицита восприятия: малого времени предъявления, низкого контраста, при наложении шума и т. п.

В отличие от предыдущих исследований [Coffin, 1978; Gervais et al., 1984] при вычислении преобразования Фурье И. А. Вол и М. Б. Павловская (Вол, Павловская, 1986) учитывали важность выбора начала отсчета, связанного с изображением. Изменение начала отсчета не влияет на амплитудный спектр изображения, а меняет только его фазовый спектр. Известно, что для реальных изображений фазовая характеристика более значима, чем амплитудная из-за схожести у реальных изображений амплитудных характеристик [Campbell, 1980]. Поэтому необходимо вводить каким-либо образом начало координат отдельно для каждого изображения. Вол и Павловская [Вол, Павловская, 1986; Vol et al., 1991] проанализировали несколько возможностей ввода начала координат для изображений. Они исходили из предпосылки минимизации фазового спектра на низких частотах. В результате в одном из вариантов минимизации получили в качестве точки начала отсчета для изображений центр тяжести.

Для моделирования в работах [Вол, Павловская, 1986; Бондарко, 1989; Vol et al., 1990] использовались экспериментальные данные, полученные на трех алфавитах: абстрактных фигурах, образованных квадратами, контурных фигурах и стимулах, подобных буквам. Были получены высокие корреляции между экспериментальными данными и модельными расчетами.

В следующей работе И. А. Вол [Вол, 1988] моделировал процесс дискриминации пространственных частот уже с помощью модели модулей. В результате у зависимости дифференциальных порогов от частот выявилось несколько локальных максимумов. Подобная зависимость наблюдалась и некоторыми экспериментаторами в психофизических и поведенческих исследованиях.

Ранее [Бондарко, 1989] для уточнения механизмов классификации изображений рассматривалось несколько моделей, оценивающих близость между изображениями по расстояниям между ними. Первая модель заключалась в том, что расстояние между объектами

оценивалось с помощью разности их спектров при совмещении центров тяжести изображений. Во второй модели расстояния вычислялись тоже как нормы от разностей спектров изображений, но при таком их совмещении, чтобы у корреляционной функции от функций освещенностей изображений был максимум. Эта модель часто рассматривается при моделировании процесса различения текстур [см. например, Caelli et al., 1986]. Ее часто называют также корреляционной моделью. Она рассматривается и как один из возможных вариантов модели согласованной фильтрации. При подсчете ответов моделей для контурных изображений использовались как непосредственно сами контуры, так и заполненные контурные изображения. В третьей модели — одной из возможных реализаций модели модулей — расстояния вычислялись как суммы разностей от конечного набора коэффициентов Фурье, полученных при разложении изображений в двумерные ряды Фурье. Следует отметить, что размер модуля подбирался таким образом, чтобы энергия отфильтрованных изображений оставалась максимальной. Оказалось, что конкретное изображение наилучшим образом описывает модуль своего размера. Для каждого изображения была выбрана точка начала отсчета таким образом, чтобы коэффициенты при синусоидальных базисных функциях, соответствующих первой гармонике, были равны нулю. Это означает, что фазовая характеристика на первых гармониках была минимизирована. В результате оказалось, что лучшее приближение ко всем экспериментальным данным дает модель, вычисляющая расстояния как норму разностей спектров при совмещении у изображений центров тяжести, а не модель модулей или модель согласованной фильтрации.

В работе Калитиевского и др. [Kaliteevsky et al., 1994] модель модулей получила свое дальнейшее развитие. В ней предлагается алгоритм описания изображений, который инвариантен к преобразованиям масштаба. Дело в том, что набор модулей дискретен по размеру, в то время как размер изображений меняется непрерывно. В таких условиях необходимо обеспечить одинаковое описание изображений, отличающихся только размером в виду существующей в зрительной системе инвариантности к изменению размера.

Алгоритм Калитиевского и др. был разработан на основе данных физиологических исследований, проведенных в Институте физиологии им. Павлова. В этих исследованиях показано, что часть рецептивных полей стриарной коры при малых контрастах изображений можно считать линейными [Glezer et al., 1980, 1982]. Ответы нейронов с такими рецептивными полями можно описать как интегралы от произведений весовых функций рецептивных полей и функций, задающих контраст стимула. Их весовые функции близки к одномерным элементам Габора [Kulikowski, Marcelja, Bishop, 1982; Marcelja, 1980]. В работе Глезера, Яковлева и Гаузельмана [Glezer, Yakovlev, Gauzelman, 1989] было показано, что распределение размеров и частотных настроек рецептивных полей стриарной коры при эксцентриситете 0–6° носит дискретный характер. В них имеются локальные максимумы, отстоящие друг от друга примерно на пол октавы. Кроме того, количество зон в профилях активности простых клеток варьирует в разных клетках от 2-х до 8-ми; то есть число циклов в весовых функциях изменяется в пределах 1–4. Распределение размеров имеет максимумы при 0,62, 0,9, 1,24, 1,8, 2,48 и 3,4 угл.град. Распределение пространственных частот образует дискретное распределение с максимумами 0,45, 0,69, 0,9, 1,35, 1,88, 2,7, 3,8 и 5,6 циклов / град. Если разделить размеры РП на размеры периодов частот, то среднее количество циклов в весовых функциях окажется равным величинам: 1, 1,41, 2,0, 2,9, 4.15.

Эти данные, а также данные по оценке ориентационной настройке частотных каналов [например, Wilson, McFarlane, Phillips, 1983] были взяты за основу модели Калитиевского и др. [Kaliteevsky et al., 1994]. В ней в одном модуле взято 10 элементов Габора (5 четных косинусов и 5 нечетных синусов) с разной частотой настройки (подобные весовые функции показаны на рис. 1). Ориентационная настройка модулей отличается на 30°. Модули разного размера и ориентации центрированы. Модули одного размера вычисляют код изображения в виде 60 чисел (10 весовых функций в 6 ориентациях). И для пространственных размеров модулей и для частот их гармоник выбраны значения логарифмических рядов с шагом в пол октавы (1.0, 1.41, 2.0, 2.83, 4). Специфичность алгоритма заключается во введении масштабно-весовой функции, которая позволяет передать

инвариантное описание изображения, подверженное непрерывным изменениям в размере.

На первом этапе изображение на входе I(x, y) обрабатывается несколькими каналами, настроенными на различные частоты. В K-том канале изображение I(x, y) на входе сначала свертывается с функцией — разностью гауссиан, а затем обрабатывается K-тым модулем. Выходной сигнал такого модуля включает информацию об обрабатываемом изображении. Поэтому самый простой способ найти код изображения — выбрать, при помощи определенного критерия, наиболее возбужденную колонку или модуль (например, номер K).

Примечательно, что в такой модели код изображения остается постоянным, только если размер изменяется дискретно в $\sqrt{2}^i$ раз, где i — целое число. Чтобы достичь большей масштабной инвариантности, было предложено выбрать группу наиболее возбужденных колонок для получения кода картинки. Выходной сигнал каждого модуля учитывается с определенным весом, меняющимся по экспоненциальному закону. Корректность работы данного алгоритма была проверена на ряде изображений. В результате симуляции были получены хорошие результаты. Был сделан вывод, что этот метод производит непрерывную интерполяцию коэффициентов, несмотря на дискретность модулей. Авторы не использовали свою модель для аппроксимации каких-либо экспериментальных данных по опознанию или различению изображений.

Модель модулей может использоваться для интерпретации различных результатов психофизических исследований. Эта модель, как следует из вышеизложенного, может неплохо объяснять данные по восприятию простых изображений, где наблюдается инвариантность к размеру. В случае опознания сложных изображений требуется вводить в рассмотрение взаимодействий между модулями, поскольку в описание таких изображений вовлекаются и низкие и высокие частоты. Природа взаимодействия между модулями не исследована и не известна. Поэтому пока преждевременно использовать их в модели.

В качестве примера приведем моделирование моделью модулей недавно полученных нами результатов по изучению иллюзии Оппель-Кундта [Бондарко и др., 2016]. Иллюзия заключается в том, что

заполненный параллельными вертикальными линиями пространственный промежуток (пространственный интервал) кажется более широким, чем промежуток, ограниченный только двумя параллельными вертикальными линиями по краям. (Заметим, что подобная иллюзия возникает при любой ориентации линий.) Была получена инвариантная по отношению к размеру промежутка зависимость величины иллюзии от числа находящихся внутри него линий. Максимум величины иллюзии приходился на семь линий. Подобная инвариантная по отношению к величине пространственных интервалов зависимость противоречит гипотезе, предложенной Булатовым и др. [Булатов и др., 1995, 1997], о возникновении иллюзии Оппель-Кундта за счет фильтрации изображений пространственно-частотными каналами с последующим сравнением размеров отфильтрованных изображений. Максимум величины иллюзии, согласно этой гипотезе, приходится на наиболее чувствительный канал, где период частоты равен расстоянию между соседними линиями. В рамках этих предположений, при увеличении величины промежутка в два раза максимум должен был сдвинуться в сторону большего количества линий. Однако согласно нашим данным этого не происходит.

Каким образом можно интерпретировать полученные нами данные?

Как и многие исследователи, мы предполагаем, что описание объектов в первом приближении может осуществляться конечным набором пространственно-частотных фильтров. Каждый объект оптимальным образом (с наименьшей потерей информации) описывается своим набором фильтров, заданных на определенном участке поля зрения. Объекты кажутся равными по размеру, если описывающие их фильтры заданы на областях одного размера. Этих положений придерживаются в своих исследованиях и Булатов с соавторами [Bulatov et al., 1997; Bulatov, Bertulis, 1999]. Разница заключается в том, что мы предполагаем объединение нескольких фильтров, настроенных на разные частоты, в модули. Модули состоят из фильтров, имеющих от одной до нескольких гармоник и обрабатывающих один и тот же участок поля зрения.

На рис. 3 показаны результаты расчетов размера промежутка в модели модулей при оптимальном описании изображений, состоя-

щих из нескольких линий, элементами Габора. Здесь модули состоят из 1–5 гармоник. За размер модуля был взят промежуток, на котором значения гауссиан, огибающих синусы и косинусы, превышают уровень 10 % от максимального значения. По оси абсцисс отложено количество линий в изображениях, по оси ординат — величина промежутка. Оказалось, что лучшее приближение к экспериментальным кривым дают модули, состоящие из четырех или пяти гармоник. При увеличении числа гармоник в модуле максимум зависимости сдвигается в сторону увеличения числа линий. Модель объясняет возникшие искажения в оценке размеров пространственных интервалов при добавлении дополнительных линий.

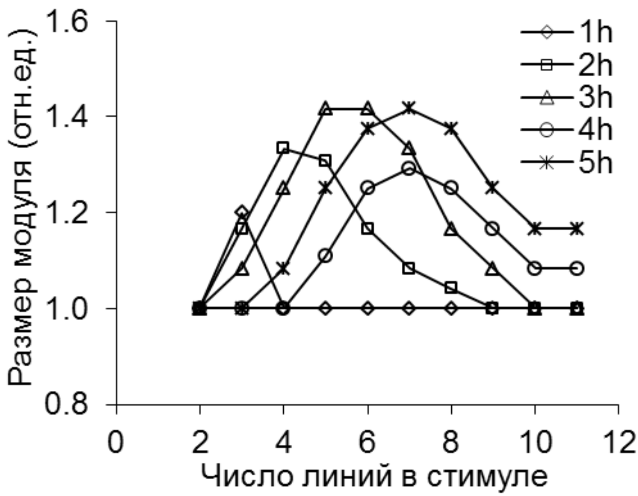


Рис. 3. Величина иллюзии Оппель-Кундта, рассчитанная в модели модулей

Ось абсцисс — количество линий в стимуле. Ось ординат — размер модуля (размер пространственного интервала), при котором происходит наилучшее описание формы стимула в относительных к размеру стимула единицах. Кривая 1h –модуль состоит из одной гармоники, кривые 2–5h- из 2–5 гармоник.

В пользу модели модулей высказываются и другие исследователи. Так, подобное объяснение данных по мгновенному определению количества изображений в поле зрения наличием фильтров с не-

сколькими периодами в первичной зрительной коре (максимально с четырьмя) дают Аткинсон и др. [Atkinson et al., 1976] и последующие исследователи этой проблемы.

В настоящее время наибольшее распространение получили многослойные модели, учитывающие взаимодействия между пространственными элементами. Взаимодействия исследуются, например, при изучении иллюзии наклона, где в качестве стимулов используют отстоящие друг от друга на различных расстояниях элементы Габора с разной ориентацией [Morgan et al., 2000]. В иллюзии наклона, как правило, выявляются тормозные связи. Напомним, что впервые эта иллюзия была объяснена Блэйкмором и др. [Blakemore et al., 1970] торможением между ориентационными каналами. В нейрофизиологии выявлено оппонентное торможение между разными ориентациями и на уровне отдельных нейронов [Шелепин, 1981а, 1981б, 1982].

Иного рода взаимодействия выявляются при решении других зрительных задач. Возбудительные связи между пространственными элементами могут обуславливать процесс группировки зрительных объектов, описанный в гештальтпсихологии принципами хорошего продолжения, пространственной близости между элементами объединения, их подобием, замкнутостью контура и др. Подробно эти принципы с современной точки зрения изложены в обзоре Вагемана и др. [Wagemans et al., 2012]. В более конкретной задаче — интерполяции фрагментов, образованных элементами Габора — Филд и др. [Field et al., 1993] показали, что расположение фрагментов вдоль гладкого контура (принцип хорошего продолжения- good continuation) улучшает способность идентифицировать контур в шуме. Они определили возможную вариабельность между ориентациями соседних элементов Габора, не препятствующую их объединению. Затем сопоставили свои результаты с известными морфофизиологическими данными по связям между ориентационными колонками в первичной зрительной коре [Hubel, Weisel, 1974a, b; Rockland, Lund, 1983; McGuire et al., 1991]. Согласованность психофизических и физиологических данных оказалась приемлемой. Известно, что интерполяция происходит для изображений достаточно больших размеров [Бондарко и др., 2010, 2016; Вахрамеева и др., 2008] и кроме того наблюдается влияние дополнительных зрительных объектов на

восприятие тестовых на больших расстояниях. Поэтому требуются дальнейшие исследования по изучению взаимодействий между рецептивными полями нейронов первичной зрительной коры и более высоких уровней зрительной системы.

Литература

Бондарко В. М. Сопоставление метрик, заданных на пространстве признаков изображений, с ошибками в опознании этих изображений // Сенсорные системы. 1989. Т. 3. N1. С. 56–68.

Бондарко В. М., Бондарко Д. В., Чихман В. Н. Зависимость иллюзии Оппель-Кундта от образующих ее линий // Экспериментальная психология. 2016. Т. 9. № 4. С. 5–12.

Бондарко В. М., Бондарко Д. В., Чихман В. Н. Влияние шума на опознание фрагментированных контурных изображений // Физиология человека. 2016. Т. 42. № 3. С. 94–102.

Бондарко В. М., Гаузельман В. Е. Перестройка пространственных элементов зрительной системы в зависимости от локального изменения освещенности // Сенсорные системы. 1987. Т. 1. N2 С. 199–204.

Бондарко В. М., Гаузельман В. Е. Пространственная организация элементов, выявляемых в условиях обнаружения или опознания зрительных стимулов // Физиология чел. 1988. Т. 14. N2. С. 204–211.

Бондарко В. М., Гаузельман В. Е., Глезер В. Д. Механизмы зрительного обнаружения и опознания различны // Физиология человека. 1983. Т. 9. N8. С. 496–498.

Бондарко В. М., Гаузельман В. Е., Чихман В. Н. Изменение контрастных порогов обнаружения узкой полосы в процессе тренировки // Сенсорные системы. 1999a. Т. 13. С. 189–194.

Бондарко В. М., Данилова М. В. Связь краудинг-эффекта с функционированием высокочастотных пространственных элементов // Сенсорные системы. 2002. Т. 16. N2. С. 89–99.

Бондарко В. М., Данилова М. В., Красильников Н. Н., Леушина Л. И., Невская А. А., Шелепин Ю. Е. Пространственное зрение. Спб.: Наука. 1999б. 218 С.

Бондарко В. М., Семенов Л. А. Острота зрения и краудинг-эффект у школьников 8–17 лет // Физиология человека. 2005. Т. 31. N5. С. 13–21.

Бондарко В. М., Семенов Л. А. Острота и гиперострота зрения у школьников 11–17 лет // Физиология человека. 2012. Т. 38. № 3. С. 56–61.

Бондарко В. М., Семенов Л. А., Солнушкин С. Д., Чихман В. Н. Инвариантность к размеру при восприятии иллюзорных и фрагментированных контуров в зависимости от возраста наблюдателей //Физиология человека. 2010. Т. 36. № 6. С. 7–14.

Булатов А., Бертулис А., Кочанене Й. Иллюзия Оппель-Кундта и анизотропия зрительного поля // Сенсорные системы. 1997. Т. 11. N. 3. С. 312–316.

Булатов А. Н., Бертулис А. В., Мицкене Л. И. Количественное исследование геометрических иллюзий // Сенсорные системы. 1995. Т. 9 (2–3). С. 79.

Вахрамеева О. А., Шелепин Ю. Е., Мезенцев А. Ю., Пронин С. В. Изучение восприятия неполных контурных изображений различного размера // Рос. Физиол. Журн. им. И. М. Сеченова. 2008. Т. 94. N10. С. 1158–169.

Вол И. А. Пространственно-частотная модель гипероcтроты зрительной системы // Сенсорные системы. 1988. Т. 2. N2. С.133–138.

Вол И. А., Павловская М. Б. Корреляция между близостью Фурье-спектров изображений и ошибками в их опознании // Физиология человека. 1986. Т. 12. N3. С. 400–406.

Глезер В. Д. Зрение и мышление. Л.: Наука, 1985. 300 С.

Глезер В. Д. Зрение и мышление СПб.: Наука, 1993. 285 с.

Глезер В. Д. Кусочный Фурье-анализ изображений и роль затылочной, височной и теменной коры в зрительном восприятии // Физиол. Журн. СССР им. И. М. Сеченова. 1978. Т. 64. С. 1719–1730.

Глезер В. Д., Гаузельман В. Е. Модули стриарной коры кошки // Сенсорные системы. 2001. Т. 15. № 1. С. 29–35.

Глезер В. Д., Яковлев В. В., Гаузельман В. Е. Рецептивные поля нейронов зрительной коры кошки: организация, классификация, функция // Сенсорные системы. 1987. Т. 1. № 4. С. 427–438.

Данилова М. В., Бондарко В. М. Влияние контекста на выполнение задачи различения ориентации прямоугольных решеток на пределе разрешения зрительной системы // Сенсорные системы. 2002. Т. 16. N2. С. 100–109.

Семенов Л. А., Чернова Н. Д., Бондарко В. М. Измерение остроты зрения и краудинг-эффекта у детей в возрасте от трех до девяти лет // Физиология человека. 2000. Т. 26. С. 67–72.

Шелепин Ю. Е. Ориентационная избирательность и пространственно-частотные характеристики рецептивных полей нейронов затылочной коры кошки. Нейрофизиология, 1981, том 13, № 3, стр. 227–232.

Шелепин Ю. Е. Фильтрационные свойства рецептивных полей нейронов зрительной коры. Доклады АН СССР, том 261, № 6,1981, стр. 1506–1509.

Шелепин Ю. Е. Пространственно-частотные характеристики и ориентационная избирательность рецептивных полей нейронов зрительной коры. Сб. «Переработка информации в зрительной системе. Высшие зрительные функции». Наука, Л., 1982, стр.28–35.

Шелепин Ю. Е., Бондарко В. М. Разрешающая способность и дискретизация изображений в зрительной системе // Рос. физиологический ж. им. И. М. Сеченова. 2002. Т. 88. N9.С. 1116–1132.

Шелепин Ю. Е., Глезер В. Д., Бондарко В. М., Павловская М. Б., Вол И. А., Данилов Ю. П. Пространственное зрение // Физиология зрения / Ред. А. Л. Бызов. М.: Наука, 1992. С. 528–586.

Atkinson J. Review of human visual development: crowding and dyslexia // Vision and visual dysfunction (ed.J.R.Cronly- Dillon). 1991. V. 13. Vision and visual dislexia (ed.J.F.Stein). P. 44–57.

Atkinson J., Campbell F., Francis M. The magic number 4 +/- 0: A new look at visual numerosity // Perception. 1976. V. 5. P. 335–342.

Blakemore C., Campbell F. W. On the existence in human visual system of neurones selectively sensitive to the orientation and size of retinal image // J. Physiology. 1969. V. 203. N1. P. 237–260.

Blakemore C., Carpenter R. H.S., Georgeson M. A., 1970. Lateral inhibition between orientation detectors in the human visual system // Nature. 1970. V. 228. P. 37–39.

Bulatov A., Bertulis A. Distortions of length perception // Biological Cybernetics. 1999. V. 80. P. 185–193.

Bulatov A., Bertulis A., Mickiene L. Geometrical illusions: Study and Modelling // Biological Cybernetics. 1997. V. 77. P. 395–406.

Campbell F. W. The physics of visual perception // Phil. Trans. R. Soc. Lond. Ser.B. 1980. V. 290. № 1038. P. 5–9.

Campbell F. W., Cooper C., Robson J., Sachs M. The spatial selectivity of visual cells of the cat and squirrel monkey // J. Physiol. 1969. V. 204. P. 120–122.

Campbell F. W., Robson J. G. Application of Fourier Analyses to the Visibility of Gratings // J. Physiol. 1968. V. 197. P. 551–566.

Coffin S. Spatial frequency analysis of block letters does not predict experimental confusions // Percept. and Psychophys. 1978. V. 23. N.1. p.69–74.

Danilova M. V., Bondarko V. M. Foveal contour interactions and crowding effects at the resolution limit of the visual system // J. Vision. 2007. V. 7(2). N. 25. P. 1–18.

Field D. J., Hayes A., Hess R. F. Contour integration by the human visual system: Evidence for a local "association field" //Vision Res. 1993. V. 33. P. 173–185.

Flom M. C. Contour interaction and the crowding effect // In R. P. Rutstein (Ed.). Problems in Optometry. 1991. V. 3. № 2. P. 237–257.

Flom M. C., Weymouth F. W., Kahneman D. Visual resolution and contour interaction // J. Opt. Soc. Am. 1963a. V. 53. P. 1026–1032.

Flom M. C., Heath G., Takahaski E. Contour interaction and visual resolution: Contralateral effect // Science. 1963b. V. 142. P. 979–980.

Gervais M. J., Harvey L. O., Roberts Y. O. Identification confusions among letters of the alphabet // J. Exp. Psychol. Hum. Percept. and Perform. 1984. V. 10. N5. P. 655–666.

Glezer V. D., Tcherbach T. A., Gauzelman V. E., Bondarko V. M. Linear and non-linear properties of simple and complex receptive fields in area 17 of the cat visual cortex: A model of the fields // Biol. Cybern. 1980. V.37. P. 195–208.

Glezer V. D., Yakovlev V. V., Gauselman V. E. Harmonic basis function for spatial coding in the cat striate cortex // Visual Neurosci. 1989. V. 3. P. 351–383.

Hubel D. H., Wiesel T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex // J. Physiol. 1962. V. 160. P. 251–260.

Hubel D. H., Weisel T. N. Sequence regularity and geometry of orientation columns in the monkey striate cortex //J. Comparative Neurology.1974a. V. 158. P. 267–294.

Hubel D. H., Weisel T. N. Uniformity of monkey striate cortex: A parallel relationship between field size, scatter and magnification factor // J. Comparative Neurology. 1874b. V. 158. P. 295–306.

Kaliteevsky N. A., Semenov V. E., Glezer V. D., Gauselman V. E. Algorithm of invariant image description by the use of a modified

Gabor transform // Applied optics. 1994. V. 33. N23. P. 5256–5261.

Kulikowski J. J., King-Smith P. E. Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation // Vision Res. 1973. V. 13. P. 1451–1463.

Kulikowski J. J., Marcelja S., Bishop P. O. Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex // Biological Cybernetics. 1982. V. 43. P. 187.

Levi D. M. Crowding — An essential bottleneck for object recognition: A mini-review //Vision research. 2008. V. 48. P. 635–654.

McGuire, B. A., Gilbert, C. D., Rivlen, P. K. & Weisel, T, N. Targets of horizontal connections in macque Primary visual cortex // J. Comparative Neurology. 1991. V. 305. P. 370–392.

Maffei L., Fiorentini A. The visual cortex as a spatial frequency analyser // Vision Research. 1973. V.13. N7. P. 1255–1267.

Maffei L., Fiorentini A. Spatial frequency rows in the striate visual cortex // Vision Research. 1977. V.17. P. 257–264.

Marcelja S. Mathematical description of the responses of simple cortical cells // J. Opt. Soc. Am. 1980. V.70. P. 1297–1300.

Morgan M. J., Mason A. J.S., Baldassi S. Are there separate first-order and second-order mechanisms for orientation discrimination? // Vision Research. 2000. V. 40. P. 1751–1763.

Movshon J.A., Thompson I. D., Tolhurst D. J. Spatial summation in the receptive field of simple cells in the cat's striate cortex // J. Physiology. 1978. V. 283. P. 53–77.

Palmer L. A., Davis T. L. Reseptive-field structure in cat striate cortex // J. Neurophysiology. 1981. V.46. P. 260–276.

Polat U., Sagi D. Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiment // Vision Research. 1993. V. 23. № 7. P. 993–999.

Sillito A. M. The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat // J. Physiology. 1975. V.250. P. 305–329.

Sillito A. M. Inhibitory mechanisms influencing complex cell orientation selectivity and their modification at high resting discharge levels // J. Physiology. 1979. V. 289. P. 33–53.

Shapley R. M., Tolhurst D. J. Edge detector in human vision // J. Physiology. 1973. V. 229. P. 165–183.

Rockland K., Lund J. S. Intrinsic laminar lattice connections in primate visual cortex // J. Comparative Neurology. 1983. V. 216. P. 303–318.

Vol I. A., Pavlovskaya M. B., Bondarko V. M. Similarity between Fourier transform of images predicts their experimental confusion // Perception a. Psychophys. 1990. V. 47. N1. P. 12–21.

Wagemans J., Elder J. H., Kubovy M., Palmer S. E. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization // Psychological Bulletin, 2012. V 138(6). P. 1172–1194.

Wilson H. R., Gelb D. J. Modified line element theory for spatial frequency and width discrimination // J. Opt. Soc. Amer. A. 1984. V. 1. 124–131.

Wilson H. R., McFarlane D.K., Phillips G. C. Spatial frequency tuning of orientation selective units estimated by oblique masking // Vision Research. 1983. V. 23. № 9. P. 873–882.

Глава 7.
Нейронные сети, обеспечивающие бинокулярное зрение

С. В. Алексеенко

Аннотация

На основе имеющихся анатомических сведений об организации зрительных путей из сетчатки и нейрофизиологических данных о функциональных свойствах нейронов первичной зрительной коры построена карта представительства во входном слое коры объектов 3-D пространства. Рассмотрены структурные особенности формирования сети диспаратно-селективных бинокулярных нейронов, которые настроены на объекты, расположенные ближе, дальше и на плоскости фиксации, объяснена природа кодирования абсолютной диспаратности в этом поле. Прослежены пути проведения информации от бинокулярных нейронов коры к мотонейронам парных черепных ядер нервов III и VI слоев, которые обеспечивают вергентные движения глаз.

Бинокулярное зрение характеризуется двумя основными феноменами: 1) слияние (фузия) двух сетчаточных изображений объекта, и 2) стереовосприятие (стереосинтез, стереопсис) — ощущение объемности формы объектов и восприятие удаленности объекта относительно наблюдателя. Слитное видение (без двоения) возникает для объектов, проецирующихся на так называемые корреспондирующие точки сетчаток. Это центральные ямки сетчаток, а также участки, равноудаленные от них в одном и том же направлении. Совокупность точек пространства, которые проецируются на корреспондирующие точки двух сетчаток, называемая гороптером, имеет в горизонтальной плоскости форму, близкую к окружности Вайс-Мюллера. Эта окружность проходит через оптические центры обоих глаз и точку бификсации. Объекты, проецирующиеся на некорреспондирующие, диспаратные, точки сетчаток, воспринимаются двоящимися (диплопия), но наблюдатель может определить, ближе или дальше плоскости фиксации они расположены.

Восприятие третьей координаты пространства основано на реконструкции в нейрональных сетях головного мозга картины окру-

жающего мира из двумерных изображений на сетчатках глаз. Начальный этап реконструкции — объединение входных сигналов из двух глаз — происходит в первичной зрительной коре. В нижележащей зрительной структуре, наружном коленчатом теле, проводящие возбудительные пути из двух сетчаток не конвергируют, а проходят раздельно через разные слои (рис.1).

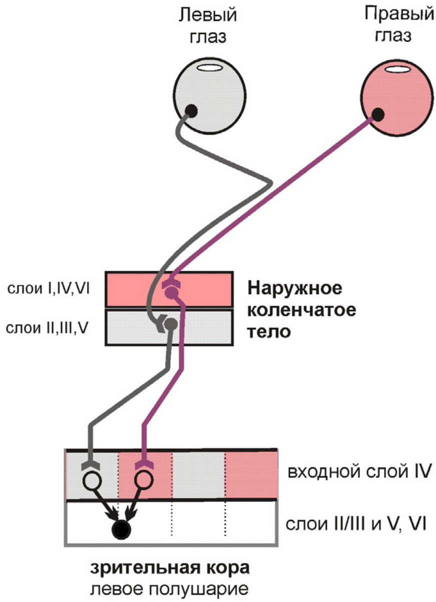


Рис. 1. Схема зрительных путей из сетчаток в зрительную кору

Однако в первичной зрительной коре V1 обнаружены клетки, которые реагируют на стимуляцию обоих глаз; такие бинокулярные нейроны чувствительны к различиям (диспаратности) между изображениями объекта на сетчатках [Hubel, Wiesel, 1962; Barlow et al., 1967; Poggio, Fischer, 1977]. К настоящему времени установлено, что чувствительность к ретинальной диспаратности обеспечивается разницей в положениях левого и правого рецептивных полей бинокулярных нейронов (позиционный механизм), а также разницей в фазовом сдвиге между профилями рецептивых полей разных глаз при совпадении положений этих полей (фазовый механизм) [Barlow

et al., 1967; Nikara et al., 1968; Freeman, Ohzawa, 1990; Ohzawa et al., 1990; DeAngelis et al., 1991; Poggio, 1995; Anzai et al., 1999].

Совокупность бинокулярных нейронов коры, избирательных к разным значениям диспаратности, сигнализирует о положении в пространстве, что необходимо для восприятия объемной формы объектов, их местоположения, а также для управления движениями глаз.

Представительство полей зрения глаз в коре головного мозга

У млекопитающих с фронтальным расположением глаз, а также у человека, информация о зрительных объектах поступает в первичную зрительную кору головного мозга по четырём раздельным путям, исходящим из назальной и темпоральной половин двух сетчаток. Каждый из этих путей обеспечивает во входном слое IV зрительной коры двумерное топографически упорядоченное отображение половины поля зрения одного глаза. Однако, вследствие частичного перекреста в оптической хиазме зрительных путей из сетчаток, в каждом полушарии представлены контралатеральные (противоположные) половины полей зрения обоих глаз, что удобно для формирования бинокулярных нейронов, имеющих рецептивные поля на обеих сетчатках (рис.2).



Рис. 2. Представительство полей зрения глаз в разных полушариях

К настоящему времени установлено, что в каждом полушарии также представлена и часть ипсилатеральной половины полей зрения

глаз. Это показано при изучении у приматов проводящих путей из сетчатки, а также в клинических и психофизических исследованиях у человека [Stone et al., 1973; Bunt et al., 1977; Leventhal et al., 1988; Fukuda et al., 1989; Fendrich et al., 1996; Reinhard et al., 2003; Marzi et al., 2009]. На уровне центрального горизонтального меридиана поля зрения ширина части ипсилатерального полуполя составляет 0.6–1.5 угл. град. И она постепенно увеличивается до 15 угл. град на крайней нижней периферии и до 5–9 угл. град. на верхней. Следовательно, центральные области полей зрения глаз по обе стороны от вертикального меридиана представлены в обоих полушариях. Источником области двойного представительства является зона назо-темпорального перекрытия на сетчатке. Часть ганглиозных клеток этой зоны посылают аксоны в одно полушарие, а часть клеток — в другое.

Карта проекции 3-D пространства во входном слое первичной зрительной коры

Бинокулярные нейроны поля V1, кодирующие положение в пространстве, формируются путем конвергенции аксонов монокулярных нейронов, получающих афферентацию из разных глаз (рис.1). Такие монокулярные нейроны локализованы во входном слое IV зрительной коры [Hubel, Wiesel, 2005]. Для понимания организации нейронных связей, необходимых для формирования бинокулярных нейронов, настроенных на разные положения в бинокулярно видимом пространстве, рассмотрим местоположение проекций точечных объектов в слое IV коры, т.е. слое, который состоит из монокулярных клеток. Ретинотопическая упорядоченность первичной коры позволяет определить местоположение проекций из левой и правой сетчаток с учетом организации ретино-геникулятных путей.

В зависимости от направленности зрительных волокон, исходящих из сетчатки, на ней можно выделить три части: назальную зону (Н), проецирующуюся в контралатеральное полушарие, темпоральную зону (Т), проецирующуюся в ипсилатеральное полушарие, и зону назо-темпорального перекрытия (Д), проецирующуюся в оба полушария. Примеры определения местоположения в слое IV коры проекций точечных стимулов приведены на рис. 3. Объект ***а*** проецируется на зону Т левой сетчатки и на зону Н правой. Посредством

перекрещивающихся в хиазме волокон из зоны Н и неперекрещивающися волокон из зоны Т объект ***а*** представлен в коре левого полушария, в ретинотопически соответствующих локусах.

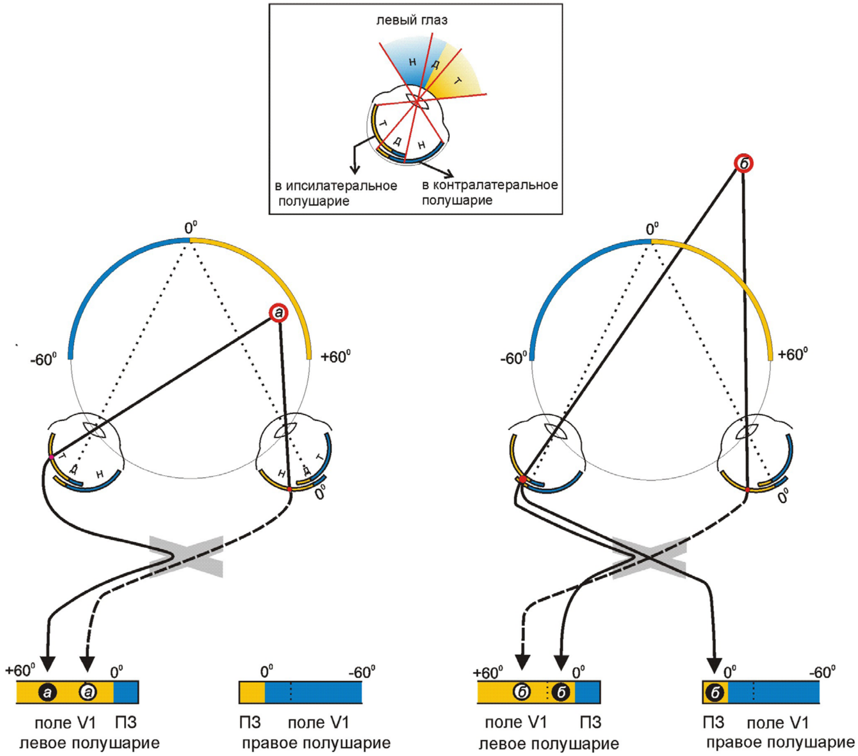


Рис. 3. Проекции объектов а и б на сетчатки глаз и их представительство в слое IV первичной зрительной коры. С целью упрощения зрительный таламус не показан

Н — назальная часть сетчатки, проецирующаяся в контралатеральное полушарие; *Т* — темпоральная часть, проецирующаяся в ипсилатеральное полушарие; *Д* — зона назотемпорального перекрытия, проецирующаяся в оба полушария; *ПЗ* — переходная зона полей *V1/V2,* где представлена часть ипсилатерального полуполя зрения. Через точку фиксации взора и оптические центры глаз проведена окружность Вайс-Мюллера. Черные и белые кружки — монокулярные локусы (нейроны), получающие вход из левого и правого глаза, соответственно. Местоположение монокулярных локусов показано на диаграммах, отображающих выпрямленную поверхность коры. Длинная ось диаграмм приблизительно соответствует проекции горизонтального меридиана поля зрения

Объект ***б*** проецируется на зону Д левой сетчатки и на зону Н правой. Посредством перекрещивающихся и неперекрещивающихся волокон из зоны Д левой сетчатки объект б представлен на территории поля коры в левом полушарии и в переходной зоне V1/V2 правого полушария, в которой представлена часть ипсилатеральной (правой) половины поля зрения. Объект б также представлен в левом полушарии посредством перекрещивающихся волокон из зоны Н правой сетчатки.

При фиксированном положении глаза, на каждую из частей сетчатки проецируются объекты, расположенные в одном из трех секторов монокулярно видимого пространства. Эти секторы показаны на вставке рис. 3. Поскольку поля зрения двух глаз значительно перекрываются, то в пространстве, видимом обоими глазами, можно выделить 9 (3х3) секторов в соответствии с количеством возможных комбинаций между тремя частями двух сетчаток (ТТ, ТД, ТН, …). На рис. 4 показано расположение секторов пространства в плоскости, проходящей через точку фиксации взора (0^o) и оптические центры глаз (окружность — теоретический гороптер Вайс-Мюллера). Положение проекций точечных объектов, расположенных в каждом из девяти секторов, приведено в нижней части рисунка. Из диаграмм видно, что точечный объект может быть представлен в двух, трех или четырех монокулярных локусах коры одного или обоих полушарий в зависимости от того, на какие части сетчаток он проецируется.

При изменении положения объекта в пределах отдельного сектора количество монокулярных локусов не меняется, изменяется только их местоположение в коре (диапазон указан на диаграммах стрелками). Чем дальше объект находится от окружности Вайс-Мюллера, тем больше расстояние между представляющими его в коре монокулярными локусами (см., например, проекции объектов 1, 2 и 3). Объекты, расположенные на окружности Вайс-Мюллера, представлены в рядом расположенных корковых монокулярных локусах левого и правого глаз. Следовательно, такие локусы находятся в одной и той же гиперколонке коры (Hubel, Wiesel, 2005). Объекты, проецирующиеся на зоны назо-темпорального перекрытия обеих сетчаток (сектор ДД, объект 4), представлены дважды, в каждом полушарии есть локусы левого и правого глаз.

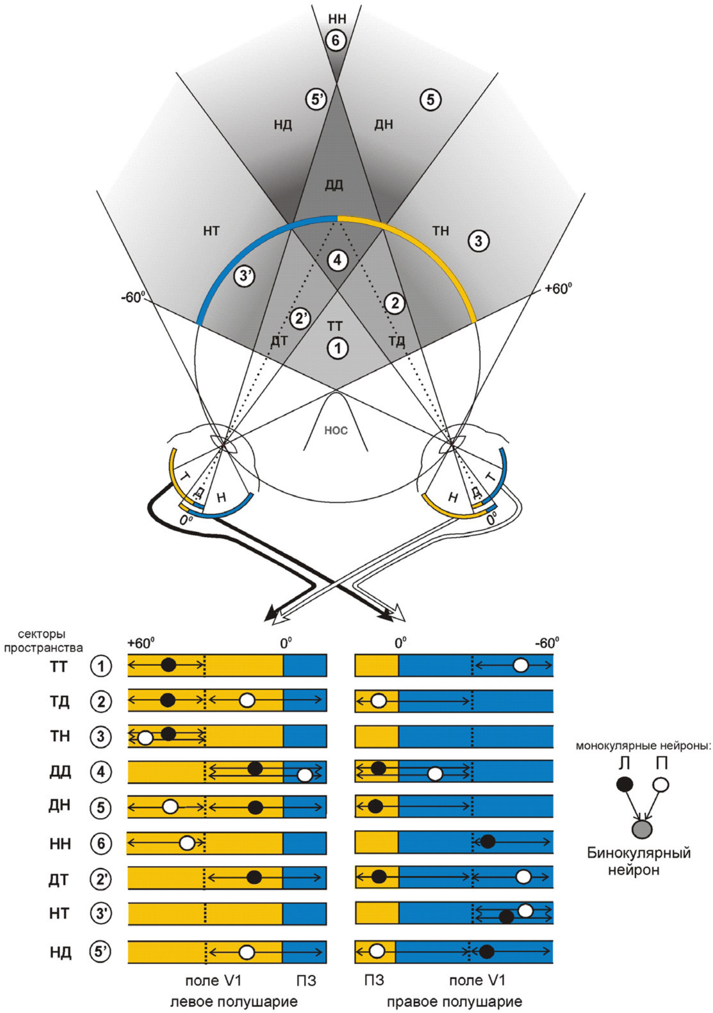


Рис. 4. Схематическое представление локализации в слое IV первичной зрительной коры проекций точечных объектов (1, 2, 2’, …), расположенных в разных секторах зрительного пространства

Местоположение проекций через левый (черный кружок) и правый (белый кружок) глаза показано на диаграммах выпрямленной поверхности коры. Длинная ось диаграмм совпадает с проекцией горизонтального меридиана поля зрения. Граница представительства в коре области Д отмечена пунктиром. Другие обозначения приведены на рис. 3

Кроме того, наблюдается еще одна закономерность в организации проекций. Для объектов, расположенных внутри и вне окружности Вайс-Мюллера, взаимное расположение левоглазного и правоглазного локусов зеркально симметрично (сравните проекции объектов 2, 4 и 3, 5 в левом полушарии).

Несмотря на то, что ширина области в полях зрения глаз, которая представлена в обоих полушариях, зависит от положения вдоль вертикального меридиана (рис. 2), надо полагать, что проекции объектов, расположенных в других горизонтальных плоскостях пространства, организованы по тем же правилам, что и в рассмотренной горизонтальной плоскости, проходящей через точку бификсации.

Таким образом, приведенная схематическая карта проекций, построенная на основе известных анатомических данных, показывает, что 1) бинокулярно видимое пространство разделяется на 9 секторов, которые по-разному представлены во входном монокулярном слое первичной зрительной коры; 2) порядок расположения этих секторов определяется только направленностью проводящих путей из назальной, темпоральной и назо-темпоральной зон двух сетчаток, поэтому он не меняется при изменении положения точки бификсации в пространстве. Карта позволяет рассмотреть особенности формирования бинокулярных нейронов в первичной зрительной коре и оценить их участие в процессе восприятия 3-ей координаты пространства и управлении движениями глаз.

Формирование бинокулярных нейронов в первичной зрительной коре

Из схематической карты расположения монокулярных проекций объектов пространства во входном слое зрительной коры (рис. 4) видно, что бинокулярные нейроны, настроенные на большинство положений пространства могут быть сформированы в противоположном полушарии при использовании внутриполушарных связей с монокулярными нейронами. В некоторых секторах пространства (ТД — объект 2, ТН — объект 3, НТ — объект 3’, ДН — объект 5) возможность формировать бинокулярные нейроны за счет внутриполушарных связей обусловлена частичным перекрестом зрительных путей из сетчатки. В других секторах (ДТ- объект 2’, ДД — объект 4,

НД — объект 5’) такая возможность появляется вследствие наличия двойных проекций из зоны назо-темпорального перекрытия (Д) на сетчатке, т.е. появления повторной проекции объекта в переходной зоне полей V1/V2, где представлена часть ипсилатерального поля зрения. Таким образом, двойные проекции обусловливают расширение области пространства, анализируемой бинокулярными нейронами, которые могут быть сформированы внутриполушарными связями. Кроме того наличие двойных проекций из зоны Д сетчатки позволяет формировать в коре обоих полушарий идентичные бинокулярные нейроны, настроенные на одно и то же положение объектов центральной области пространства (объект 4). Это увеличивает надежность оценки их положения.

Однако, бинокулярные нейроны, настроенные на положение объектов в центральных секторах пространства (ДД, НД, ДН, ТД, ДТ), могут быть сформированы не только связями с монокулярными нейронами, расположенными в том же полушарии, но также и в противоположном. Это частично подтверждают экспериментальные данные на животных с перерезкой оптической хиазмы и/или мозолистого тела [Lepore, Guillemot, 1982; Lepore et al., 1992]. Показано, что межполушарные (каллозальные) входы имеет треть бинокулярных нейронов, локализованных вблизи проекции вертикального меридиана, а также нейроны, настроенные на положения за пределами плоскости фиксации (нейроны ДАЛИ).

Однако для формирования бинокулярных нейронов, настроенных на положение объектов в секторах НН и ТТ (объекты 1, 6), необходимы межполушарные связи с монокулярными нейронами периферических участков коры. Поскольку каллозальные клетки в этих участках коры не выявлены [Innocenti, 1986], то следует заключить, что в бинокулярно видимом пространстве есть области, положение объектов в которых не может быть оценено бинокулярными нейронами первичной зрительной коры. Локализация объектов в этих секторах должна определяться с использованием целого ряда монокулярных и когнитивных признаков [Рожкова, Матвеев, 2007]. Необходимо отметить, что не только особенности локализации каллозальных клеток в коре, но и протяженность внутризональных горизонтальных нейронных связей может ограничить диапазон

диспаратностей, для выделения которых могут быть сформированы бинокулярные нейроны.

Роль бинокулярных нейронов в процессах стереовосприятия

Схема проекций точечных объектов в слой IV коры указывает, что диспаратность объектов, расположенных в плоскости фиксации, равна нулю. Бинокулярные нейроны, настроенные на объекты в этой плоскости, могут быть сформированы с использованием наиболее коротких связей с монокулярными нейронами, которые расположены в одной и той же гиперколонке коры. Размер корковой гиперколонки не превышает 1 мм [Hubel, Wiesel, 2005]. Пространственное совпадение рецептивных полей левого и правого глаз у этих нейронов объясняет почему такие объекты не воспринимаются двоящимися. Известно, что слитное видение наблюдается не только для объектов в плоскости фиксации, но и для объектов, расположенных в фузионной зоне Панума. Объекты, расположенные за пределами этой зоны, вызывают двоение (диплопию) [Ogle, 1952]. В естественных условиях наблюдения одно из «двоящихся» изображений не воспринимается, что, вероятно, обусловлено торможением со стороны структур внимания. Однако имеются доказательства того, что в таких «диплопических» случаях, в первичной коре продолжается обработка зрительной информации, обеспечивая сигналы, необходимые для управления движениями глаз и восприятия глубины.

Известно, что диспаратно-селективные нейроны, обнаружены во многих корковых полях. Однако, бинокулярные нейроны первичной зрительной коры, сформированные простой конвергенцией монокулярных клеток, могут оценить положение объектов только относительно точки бификсации, независимо от ее положения в 3-Д пространстве и относительно других объектов в зрительной сцене. Это очевидно из анализа схематической карты представительства зрительного пространства в слое IV, на которой показаны входные клетки для таких бинокулярных нейронов (локусы на рис. 4). Количество и взаимное расположение в коре активированных входов из левого и правого глаз для объектов каждого сектора пространства зависит только от частей Н, Т, Д сетчаток, на которые эти объекты проецируются. Поэтому направление взора и расстояние до точки бификса-

ции не могут изменить этот паттерн проекций из сетчаток в слой IV коры. Соответственно, при любых движениях глаз также не должна меняться и настройка диспаратно-селективных нейронов, сформированных конвергенцией монокулярных нейронов, которые расположены на определенных расстояниях в слое IV. Этот вывод, основанный только на результатах анализа проекций точечных объектов в поле V1, согласуется с данными нейрофизиологических исследований, которые указывают, что только абсолютная диспаратность кодируется бинокулярными нейронами поля V1 [Cumming, Parker, 1999; Roe et al., 2007]. На основании данных этих экспериментальных работ можно также предположить, что последующий уровень возможной конвергенции в поле V1 (конвергенция аксонов бинокулярных нейронов, сформированных конвергенцией монокулярных нейронов) не может обеспечить настройку на относительную диспаратность. Такое свойство было обнаружено у бинокулярных нейронов полей более высоких уровней, относящихся преимущественно к вентральному пути (Thomas et al., 2002; Umeda et al., 2007).

Информация об абсолютной диспаратности, обеспечиваемая бинокулярными нейронами поля V1, может быть использована для восприятия 3-ей координаты пространства и для инициации вергентных движений глаз. Многочисленные экспериментальные данные указывают, что в этом корковом поле осуществляется только предварительный анализ сигналов, необходимых для бинокулярного восприятия трехмерной формы объектов. Диспаратно-селективные нейроны поля V1 реагируют только на локальные соответствия в двух зрительных изображениях. Это было показано при использовании случайно-точечных стереограмм с противоположным знаком контраста на разных сетчатках, а также других стимулов, которые не вызывают восприятие глубины. Однако при предъявлении этих специфических стимулов возникают вергентные движения глаз, которые смещают изображения в центры фовеа [Erkelens, Collewijn, 1985; Masson et al., 1997; Cumming, Parker, 1997, 2000; Miles, 1998; Busettini et al., 2001; Erkelens, 2001].

Краткие сведения о типах движений глаз. Координированные макродвижения двух глаз при восприятии динамических зрительных сцен разделяются на 2 типа: версионные и вергентные. Верзи-

онные движения глаз осуществляют совместный поворот обоих глаз в одном и том же направлении с одинаковой амплитудой. Такие движения глаз изменяют направление взора без изменения расстояния до плоскости фиксации. Вергентные движения глаз (конвергентные и дивергентные) осуществляют поворот глаз в противоположных направлениях с одинаковой амплитудой. Такие движения глаз сдвигают положение точки фиксации для наведения обоих сетчаточных изображений объекта в центральные ямки.

Горизонтальные компоненты движений глаз обеспечиваются работой прямых латеральной и медиальной мышц глазного яблока. Управляющие сигналы к этим мышцам поступают из двух парных черепных ядер, III и VI, локализованных по обеим сторонам мозга. Прямые связи между этими ядрами и соответствующими ипсилатеральными мышцами используются для вергентных движений глаз (рис. 5). Однако для выполнения версионных движений глаз требуется другая комбинация активированных мышц, поэтому используются дополнительные непрямые (через ядро III) связи ядер VI с контралатеральными прямыми медиальными мышцами. Таким образом, команда из левого ядра VI, вызывающая сокращение мышц, обеспечивает поворот взора налево, а команда из правого ядра VI обеспечивает поворот направо.

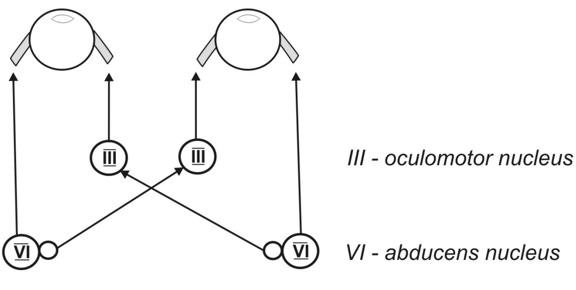


Рис. 5. Упрощенная схема иннервации прямых латеральной и медиальной мышц глазного яблока из ядер черепных нервов.

Участие нейронов коры в управлении версионными движениями глаз. Для управления версионными движениями глаз информация о глубине не требуется, нужны только сведения об угловых

координаты объекта в двумерном поле зрения, которые поступают в глазодвигательную систему из нижележащих подкорковых структур, в основном, из верхнего двухолмия [Scudder et al., 1996; Moschovakis et al., 1998]. При этом, произвольные версионные движения осуществляются при участии нейронов корковых полей. Однако в бинокулярно видимом пространстве, монокулярные оценки положения объекта через левый и правый глаза значительно различаются, если объект удален от плоскости фиксации. В связи с этим, надо полагать, что монокулярные нейроны только периферической части коры, в которой представлен монокулярный сектор пространства, дают однозначную, необходимую для версионных движений глаз, оценку направления, вдоль которого расположен объект. Информация, необходимая для управления версионными движениями глаз, может быть получена и от бинокулярных нейронов поля V1. По двумерным координатам локализации этих клеток в ретинотопически упорядоченной коре можно примерно определить направление, по которому расположен объект. Однако, экспериментально было показано, что сигналы от диспаратно-селективных нейронов коры влияют только на амплитуду версионных ответов [Masson et al., 2001].

Участие нейронов коры в управлении вергентными движениями глаз. Для управления конвергентными и дивергентными движениями глаз необходима оценка глубины положения объекта. Эти движения важны для реализации процесса бинокулярной фузии. Сигналы бинокулярных нейронов поля V1, которые настроены на положения объектов за пределами фузионной зоны Панума, могут использоваться глазодвигательной системой для инициации таких движений глаз. По терминологии, введенной Поджио Г. Ф. [Poggio, Fischer, 1977], это нейроны БЛИЗИ и нейроны ДАЛИ. Следует отметить, что большинство положений объектов, на которые настроены эти нейроны, может быть отнесено к «диплопическим», но двоение этих объектов не воспринимается.

Нейроны БЛИЗИ и ДАЛИ различаются пространственным расположением в коре входных монокулярных нейронов, кроме того, они посылают сигналы в разные глазодвигательные ядра.

Для нейронов БЛИЗИ, входные монокулярные нейроны, управляемые из ипсилатерального глаза, расположены дальше от проек-

ции центрального вертикального меридиана поля зрения по сравнению с входными нейронами контралатерального глаза (рис. 4; см. проекции объектов 2, 4 в левом полушарии и 2', 3', 4 в правом полушарии). Для нейронов ДАЛИ — наоборот: входные монокулярные нейроны, управляемые из контралатерального глаза расположены дальше (см. проекции объектов 3, 5 в левом полушарии и 5' в правом полушарии).

По данным нейрофизиологических исследований у нейронов БЛИЗИ обнаружена тенденция к доминированию входа из ипсилатерального рецептивного поля, а у нейронов ДАЛИ — из контралатерального [Poggio, Fischer, 1977; Fischer, Krueger, 1979; Fischer, Poggio, 1979; Maske et al., 1986; LeVay, Voigt, 1988; Lepore et al., 1992; Gonzalez et al., 2001; Read, Cumming, 2004]. Таким образом, более периферический монокулярный вход доминирует у обоих типов бинокулярных нейронов — перекрестный вход у нейронов ДАЛИ и неперекрестный вход у нейронов БЛИЗИ. Отметим, что перекрестные (в оптической хиазме) проводящие пути из сетчаток эволюционно более древние, они преобладают у животных с панорамным зрением.

Активация нейронов БЛИЗИ сигнализирует о положении объектов, расположенных ближе бифовеальной точки фиксации (конвергентная диспаратность). В случае необходимости перевода взора на такой объект, т.е. приближения точки бификсации к наблюдателю, сигналы нейронов БЛИЗИ передаются к мотонейронам ядер глазодвигательного нерва III, что приводит к сокращению прямых медиальных мышц обоих глаз, т.е. конвергентным движениям глаз.

Активация нейронов ДАЛИ сигнализирует о положении объектов, расположенных за пределами точки бификсации (дивергентная диспаратность). В случае необходимости перевода взора на такой объект, сигналы этих нейронов ДАЛИ поступают к мотонейронам ядер отводящего нерва VI, что приводит к сокращению прямых латеральных мышц обоих глаз, т.е. дивергентным движениям глаз.

Детальная архитектура нейронных сетей, по которым передается информация от нейронови БЛИЗИ, ДАЛИ поля V1 к мотонейронам ядер черепных нервов III и VI, до сих пор неясна. Главным образом не имеется четких данных о контактах этих функционально различных групп диспаратных нейронов с отдельными группами клеток

подкорковых центров (ростральная часть верхнего двухолмия, ретикулярная формация среднего мозга), которые реагируют при конвергентных или дивергентных движениях глаз [Mays, 1984; Judge, Cumming, 1986; Van Horn et al., 2013]. Тем не менее, имеющиеся анатомические и физиологические данные указывают, что основные пути из первичной зрительной коры к мотонейронам, которые обеспечивают вергентные движения глаз, латерализованы. Информация от диспаратно-селективных нейронов коры каждого полушария поступает в ростральную часть ипсилатерального верхнего двухолмия, затем в ипсилатеральную ретикулярную формацию, которая имеет прямые контакты с ипсилатеральными ядрами III или VI, мотонейроны которых вызывают сокращение прямой медиальной или латеральной мышц ипсилатерального глаза [Kawamura et al., 1974; Harting, 1977; Zhang et al., 1991].

Возникает вопрос, каким образом обеспечивается совместное сокращение обеих медиальных или обеих латеральных мышц глаз? Для синхронного сведения или разведения зрительных осей обоих глаз могут быть использованы сигналы, передаваемые из ретинальных зон назо-темпорального перекрытия. Объекты, которые проецируются на зоны назо-темпорального перекрытия обеих сетчаток, представлены в слое IV коры каждого полушария монокулярными нейронами, получающими входы из разных глаз (например, объект 4). На основе этих проекций в коре обоих полушарий формируются идентичные диспаратно-селективные нейроны, которые по ипсилатеральным путям синхронно передают одинаковые сигналы в верхнее двухолмие, ретикулярную формацию среднего мозга, глазодвигательные ядра III или VI, в результате чего происходит одновременное сокращение медиальных или латеральных мышц обоих глаз.

Литература

Рожкова Г. И., Матвеев С. Г. Зрение детей: проблемы оценки и функциональной коррекции. М.: Наука. 2007. 315с.

Anzai A., Ohzawa I., Freeman R. D. Neural mechanisms for processing binocular information. II.Complex cells. J. Neurophysiol. 1999. Vol. 82. P. 909–924.

Barlow H. B., Blakemore C., Pettigrew J. D. The neural mechanisms of binocular depth discrimination. J. Physiol. 1967. Vol. 193. P. 327–342.

Bunt A. H., Minckler D. S., Johanson G. W. Demonstration of bilateral projection of the central retina of the monkey with horseradish peroxidase neuronography. J. Comp. Neurol. 1977. Vol. 171. № 4. P. 619–630.

Busettini C., FitzGibbon E. J., Miles F. A. Short-latency disparity vergence in humans. J. Neurophysiol. 2001. Vol. 85. P. 1129–1152.

Cumming B. G., Parker A. J. Responses of primary visual cortical neurons to binocular disparity without depth perception. Nature. 1997. Vol. 389. P. 280–283.

Cumming B. G., Parker A. J. Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. J. Neurosci. 1999. Vol. 19, P. 5602–5618.

Cumming B. G., Parker A. J. Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the macaque. J. Neurosci. 2000. Vol. 20.P. 4758–4767.

DeAngelis G.C., Ohzawa I., Freeman R. D. Depth is encoded in the visual cortex by a specialized receptive field structure. Nature. 1991. Vol. 352. № 6331. P. 156–159.

Erkelens C. J. Organisation of signals involved in binocular perception and vergence control. Vision Res. 2001. Vol. 41. P. 3497–3503.

Erkelens C. J., Collewijn H. Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. Vision Res. 1985. Vol. 25(11). P. 1689–1700.

Fendrich R, Wessinger C. M., Gazzaniga M. S. Nasotemporal overlap at the retinal vertical meridian: investigations with a callosotomy patient. Neuropsychologia. 1996. Vol. 34. № 7. P. 637–646.

Fischer B., Krueger J. Disparity tuning and binocularity of single neurons in cat visual cortex. Exp. Brain Res. 1979. Vol. 35, P. 1–8.

Fischer B., Poggio G.F. Depth sensitivity of binocular cortical neurons of behaving monkeys. Proc. R. Soc. B: Biol. Sci. 1979. Vol. 204. P. 409–414.

Fukuda Y., Sawai H., Watanabe M., Wakakuwa K., Morigiwa K. Nasotemporal overlap of crossed and uncrossed retinal ganglion cell projections in the Japanese monkey (Macaca fuscata). J. Neurosci. 1989. Vol. 9. № 9. С. 2353–2373.

Gonzalez F., Perez R., Justo M.S., Ulibarrena C. Binocular interaction and sensitivity to horizontal disparity in visual cortex in the awake monkey. Intern. J. Neurosci. 2001. Vol. 107. P. 147–160.

Harting J.K. Descending pathways from the superior collicullus: an autoradiographic analysis in the rhesus monkey (Macaca mulatta). J. Comp. Neurol. 1977. Vol. 173. № 3. P. 583–612.

Hubel D.H., Wiesel T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol (London). 1962. Vol. 160. P. 106–154.

Hubel D.H., Wiesel T.N. Brain and visual perception. 2005. New York: Oxford UP.

Innocenti G.M. General organization of callosal connections in the cerebral cortex. In: Cerebral Cortex (eds. Jones E.G., Peters A.) New York: Plenum. 1986. Vol. 5. P. 291–351.

Judge S.J., Cumming B.G. Neurons in the monkey midbrain with activity related to vergence eye movement and accommodation. J. Neurophysiol. 1986. Vol. 55. № 5. P. 915–930.

Kawamura K., Brodal A., Hoddevik G. The projection of the superior colliculus onto the reticular formation of the brain stem. An experimental anatomical study in the cat. Exp. Brain Res. 1974. Vol. 19. P. 1–19.

Lepore F., Guillemot J.P. Visual receptive field properties of cells innervated through the corpus callosum in the cat. Exp. Brain Res. 1982. v. 46, № 3. P. 413–424

Lepore F., Samson A., Paradis M–C., Ptito M., Guillemot J-P. Binocular interaction and disparity coding at the 17–18 border: contribution of the corpus callosum. Exp. Brain Res. 1992. Vol. 90, P. 129–140.

LeVay S., Voigt T. Ocular dominance and disparity coding in cat visual cortex. Visual Neurosci. 1988. Vol. 1**, P.** 395–414.

Leventhal A. G., Ault S. J., Vitek D. J. The nasotemporal division in primate retina: the neural bases of macular sparing and splitting. Science. 1988. Vol. 240. № 4848. P. 66–67.

Marzi C. A., Mancini F., Sperandio I., Savazzi S. Evidence of midline retinal nasotemporal overlap in healthy humans: A model for foveal sparing in hemianopia? Neuropsychologia. 2009. Vol. 47. № 13. P. 3007–3011.

Maske R., Yamane S., Bishop P. O. Stereoscopic mechanisms: binocular responses of the striate cells of cats to moving light and dark bars. Proc. R. Soc. B: Biol. Sci. 1986. Vol. 229. P. 227–256.

Masson G. S., Busettini C., Miles F. A. Vergence eye movements in response to binocular disparity without depth perception. Nature. 1997. Vol. 389. P. 283–286.

Masson G. S., Busettini C., Yang D. S., Miles F. A. Short-latency ocular following in humans: sensitivity to binocular disparity. Vision Res. 2001. Vol. 41. P. 3371–3387.

Mays L. E. Neural control of vergence eye movements: convergence and divergence neurons in the midbrain. J. Neurophysiol. 1984. Vol. 51. P. 1091–1108.

Miles F. A. The neural processing of 3-D visual information: evidence from eye movements. Eur. J. Neurosci. 1998. Vol. 10. P. 811–822.

Moschovakis A. K., Kitama T., Dalezios Y., Petit J., Brandi A. M., Grantyn A. A. An anatomical substrate for the spatiotemporal transformation. J. Neurosci. 1998. Vol. 18(23), 10219–10229.

Nikara T., Bishop P.O., Pettigrew J.D. Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. Exp. Brain Res. 1968. Vol. 6, № 3. P. 353–372.

Ogle K.N. Disparity limits of stereopsis. Arch. Ophthal. 1952. Vol. 48. P. 50–60.

Poggio G.F. Mechanisms of stereopsis in monkey visual cortex. Cereb. Cortex. 1995. № 5. P. 193–204.

Poggio G.F., Fischer B. Binocular interaction and depth sensitivity in the striate and prestriate cortex of the behaving monkey. J. Neurophysiol. 1977. Vol. 40. P. 1392–1405.

Read J.C., Cumming B.G. Ocular dominance predicts neither strength nor class of disparity selectivity with random-dot stimuli in primate V1. J. Neurophysiol. 2004. Vol. 91. P. 1271–1281.

Reinhard J., Trauzettel-Klosinski S. Nasotemporal overlap of retinal ganglion cells in humans: a functional study. Invest. Ophthalmol. Vis. Sci.— 2003. Vol. 44. № 4. P. 1568–1572.

Roe, A. R., Parker A.J., Born R.T., DeAngelis G.C. Disparity channels in early vision. J. Neurosci. 2007. Vol. 27. P. 11820–11831.

Scudder C.A., Moschovakis A.K., Karabelas A.B., Highstein S.M. Anatomy and physiology of saccadic long-lead burst neurons recorded in the alert squirrel monkey. II. Pontine neurons. J. Neurophysiol. 1996. Vol. 76. № 1.P. 353–370.

Stone J., Leicester J., Sherman S.M. The naso-temporal division of the monkey’s retina. J. Comp. Neurol. 1973. Vol. 150. № 3. P. 333–348.

Thomas O.M., Cumming B.G., Parker A.J. A specialization for relative disparity in V2. Nature Neurosci. 2002. Vol. 5. № 5. P. 472–478.

Umeda K., Tanabe S., Fujita I. Representation of stereoscopic depth based on relative disparity in macaque area V4. J. Neurophysiol. 2007. Vol. 98. P. 241–252.

Van Horn M. R., Waitzman D. M., Cullen K. Evergence neurons identified in the rostral superior colliculus code smooth eye movements in 3D space. J. Neurosci. 2013. Vol. 33. P. 7274–7284.

Zhang Y., Gamlin P. D., Mays L. E. Antidromic identification of midbrain near response cells projecting to the oculomotor nucleus. Exp. Brain Res. 1991. Vol. 84. № 3. P. 525–528.

Глава 8.
Перестройка крупномасштабной нейронной сети головного мозга человека при распознавании лиц

О. В. Жукова, Ю. Е. Шелепин, Г. Е. Труфанов, В. А. Фокин, П. П. Васильев, А. В. Соколов

Введение

Крупномасштабные нейронные сети головного мозга человека объект подробного изучения последних 140 лет. Особое внимание уделяют функциональной специфике работы таких сетей при восприятии определенных классов изображений. Цель нашего исследования посвящена изучению закономерностей перестройки крупномасштабных нейронных сетей головного мозга человека, в процессе распознавания лица человека, его мимики.

В качестве актуальной модели выбрана крупномасштабная нейронная сеть головного мозга человека при восприятии лиц как важнейшего из средств невербальной коммуникации [Жукова и др., 2015; Шелепин, Жукова, 2016; Шелепин и др., 2014, 2015]. Особое значение данное направление приобретает в клинике, например, при изучении заболеваний, в ходе которых развиваются нарушения социальных когнитивных функции, в том числе в области распознавания мимики другого человека.

Методика. В качестве основного метода исследований была выбрана функциональная магнитно-резонансная томография, которая позволяет определять устойчивые паттерны активации и грубую синхронность срабатывания разных областей мозга. В наших исследованиях в качестве стимулов использованы оптоклоны виртуальных людей. Выбор данного класса изображений был обусловлен возможностью контролировать различные экспериментальные условия, например, степень выраженности эмоциональных состояний и угла поворота оптоклонов. Всего было синтезировано 36 разных лиц. В первой серии исследовании процедура стимуляции включала предъявление четырех фаз стимуляции. В фазе 1 и 3 многократно предъявляли изображения 36 оптоклонов с нейтральным выражени-

ем лица, ориентированных прямо. В фазе 2 и 4 — предъявляли изображения тех же оптоклонов, но изображения лиц давались в разных поворотах, при разных выражениях эмоции. Испытуемые добровольцы выполняли две инструкции: в фазах 1 и 2 задача состояла в определении поворота лица (направо-налево), а в фазах 3 и 4 — мимики (радость-грусть). Выбор регистрировали по нажатию испытуемыемыми левой или правой клавиши мыши. Необходимо отметить, что испытуемые не были уведомлены о наличии нейтральных прямо ориентированных лиц. Данный стимульный материал воспринимался как лица с «незаметной для глаза» минимальной степенью выраженности изучаемого признака. На рис. 1 приведены примеры изображений, предъявляемых в различных фазах стимуляции.

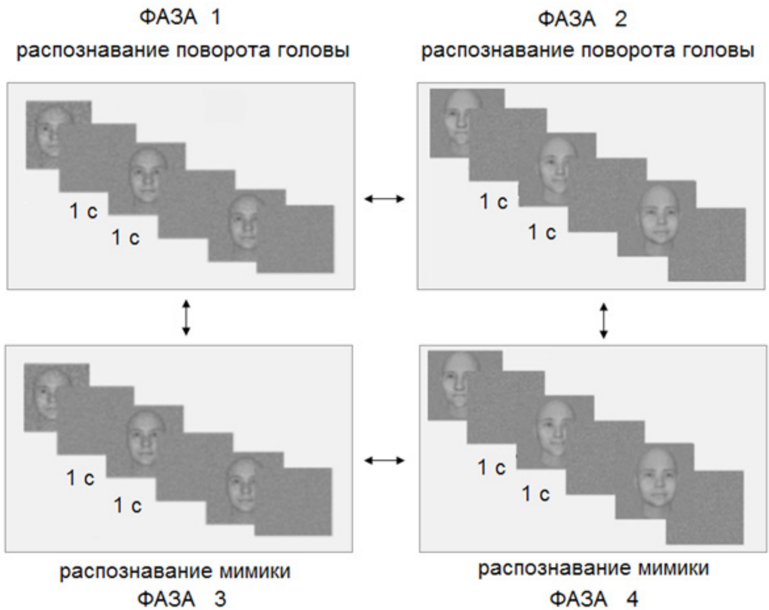


Рис. 1. Схема стимуляции: фаза 1 — предъявление нейтральных прямо ориентированных лиц в задачах определения поворота; фаза 2 — предъявление лиц, выражающих разные эмоции, ориентированные в разные стороны в задачах определения поворота; фаза 3 — предъявление нейтральных прямо ориентированных лиц в задачах определения эмоции; фаза 4 — предъявление лиц, выражающих разные эмоции, ориентированные в разные стороны в задачах определения эмоции

Результаты. Рассмотрим карты активации, полученные путем вычитания результатов в различных фазах стимуляции.

Первая рассматриваемая экспериментальная парадигма — это сравнение двух условий (фаз стимуляции), при которых инструкция оставалась одной и той же, но менялся тип изображения: предъявляли либо нейтральные прямо ориентированные лица, либо оптоклонов с мимической экспрессией, повернутых в разные стороны (рис. 2). Таким образом, менялась структура изображения, но не инструкция.

В задаче определения поворота (фаза 1- фаза 2) мы наблюдаем два паттерна активации, обозначенные на рисунке разными цветами. Красным цветом отмечена активация в условии вычитания фаза 1 — фаза 2, а голубым — обратное вычитание: фаза 2 — фаза 1. В первом случае мы наблюдаем специфический паттерн сети распознавания лиц: активация классических, ответственных за распознавание изображений лиц (например, фузиформная извилина, BA37), и дополнительных зон мозга, которые зависят от инструкции и других экспериментальных условий. В наших условиях стимуляции это области во фронтальной (BA10, BA9, BA8, BA32), теменной (BA40, BA7) и премоторной (BA6) коре.

В целом, следует отметить, что полученные карты активации зон, входящих в расширенную часть нейронной сети совпадают с данными полученными в других лабораториях в различных задачах, связанных с принятием решения [Keuken et al., 2014]. Но, в то же время, необходимо подчеркнуть, что эти же зоны играют существенную роль и в процессах восприятия лица в связи с тем, что во многих из них были найдены «зеркальные» нейроны [Rizzolatti et al., 2004]. Первоначально они были обнаружены в нижней лобной извилине, затем в нижней теменной доле и верхней височной борозде. Впоследствии в данную сеть был включен регион премоторной коры. Принято считать, что система «зеркальных нейронов» в первую очередь связана с сенсорным невербальным взаимодействием и участвует в построении мимики, позы и жестов в основном при взаимодействии субъектов. Однако существует другое предположение, что данная функция — лишь верхушка айсберга. В действительности, наблюдаемый многими отклик зеркальных нейронов, является функцией обучение нейронной сети на основе перепроверки правильности принятого решения [Шелепин, Жукова, 2016].

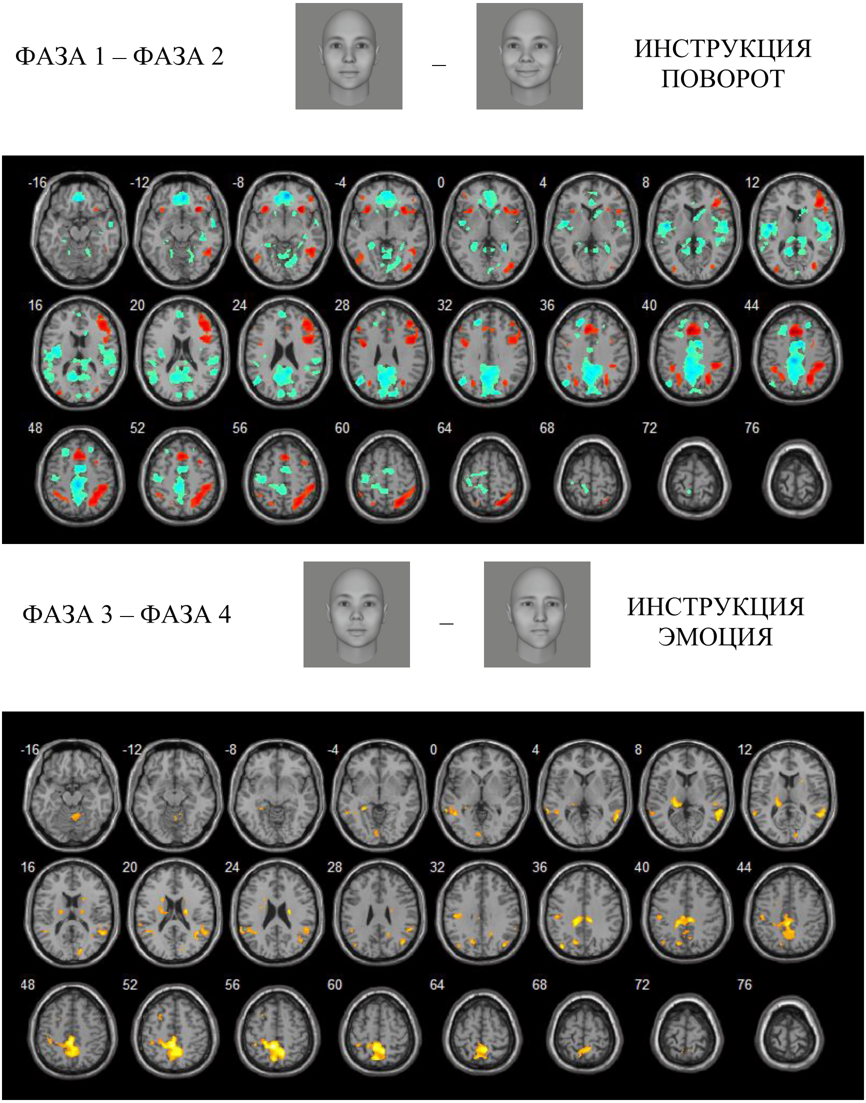


Рис. 2. Карты активации, полученные при вычитании двух фаз стимуляции: в задаче определения поворота (фаза 1-фаза 2) и мимической экспрессии лица (фаза 3-фаза 4). Данные усреднены на группе добровольцев, N=24 (без FDR-corr., $p < 0.001$)

При обратном вычитании фаз стимуляции (фаза 2 — фаза 1), мы наблюдаем отличный паттерн активации, который включает медиальную префронтальную кору (BA9, BA10, BA11, BA24), височно-теменной стык (BA39, BA40, BA22), предклинье и заднюю поясную извилину (BA23, BA29, BA30, BA31). В современных исследованиях данный паттерн определяют как «сеть по умолчанию» или «сеть базового режима» [Raichle et al., 2001]. Основная особенность данной сети заключается в том, что она наиболее активна при отсутствии внешних раздражителей, в состоянии относительного покоя. В наших экспериментальных условиях покоя не было, но стимулы, предъявляемые в фазе 2 (оптоклоны с выраженной мимической экспрессией и поворотом) казались добровольцам более очевидной, а значит, можно предположить, и более простой задачей в сравнении с фазой 1, когда предъявлялись нейтральные прямо ориентированные оптоклоны. Напомним, что добровольцы не были уведомлены об отсутствии изучаемых признаков (поворота и эмоции) в изображениях, что делало задачу крайне сложной в решении. Подтверждением данного предположения могут служить данные по ответам и времени реакции добровольцев, представленные на рис. 3–4.

Как видно из представленных данных, распознавание эмоции и поворота в условиях предъявления изображений нейтральных прямо ориентированных лиц (фазы 1 и 3) представляло значительную сложность для добровольцев, что нашло отражение в более замедленной скорости реакции. Процент нажатий добровольцев на правую и левую клавиши мыши при выполнении разных инструкций в целом располагается возле планки 50 %. Распознавание эмоции и поворота в условиях предъявления оптоклонов с выраженной мимической экспрессией и поворотом (фазы 2 и 4) оказалось более очевидной и, вероятно, легкой задачей. Так, процент правильных ответов при распознавании поворота в среднем составил 95,13 %, а эмоции — 83,26 %. Скорость реакции же, при распознавании поворота, в среднем составила 558,55 мс, а эмоции — 715,98 мс. При этом, необходимо отметить, что распознавание поворота в фазе 2 значимо отличается от распознавания эмоции в фазе 4 по всем указанным показателям ($p<0,001$).

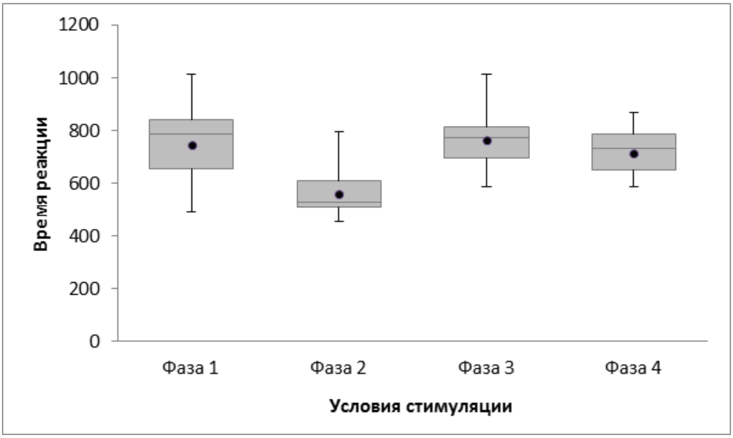


Рис. 3. Диаграмма размаха по времени реакции добровольцев при распознавании поворота и эмоции в 4 фазах стимуляции (N-22). По оси ординат указаны отклонения данного параметра от среднего для испытуемого значения. По оси абцисс — фазы стимуляции. Черные точки — среднее значение по выборке

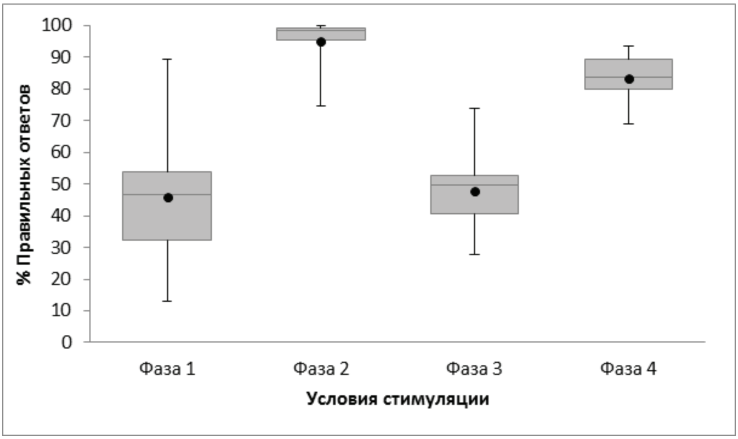


Рис. 4. Диаграмма размаха по ответам добровольцев при распознавании поворота и эмоции в 4 фазах стимуляции (N-22). По оси ординат указаны отклонения данного параметра от среднего для испытуемого значения. По оси абцисс — фазы стимуляции. В фазе 1 и 3 считали процент нажатия на правую и левую клавиши мыши. В качестве примера представлены значения по выбору правого поворота и эмоции радости. В фазах 2 и 4 — процент правильных ответов добровольцев. Черные точки — среднее значение по выборке

При анализе паттернов активации представляет интерес активность в области медиальной префронтальной коры. Как представлено на рис. 2, мы видим активацию данной области и при прямом и при обратном вычитании. Роль медиальной префронтальной коры на данный момент активно обсуждается. Так, принято считать, что данная структура участвует в описании и оценке как своего, так и чужого психического состояния [Bryan et al., 2012]. Особенно интересны исследования, изучающие роль данной области в этиологии клинических расстройств. Так, при заболеваниях, сопровождающихся дефицитом способности понимать чужие убеждения, намерения и чувства, например, при аутизме, шизофрении, тревожных расстройствах и депрессии, присутствуют дисфункции в зоне медиальной префронтальной коры. Именно эта способность часто нарушается у больных с данными заболеваниями. Так, активация данной области в составе сети базового режима, вероятно, связана с более легкой задачей в фазе 2 в сравнении с фазой 1. А активация в составе нейронной сети распознавания лиц, вероятно, связано с тем, несмотря на инструкцию определения поворота, добровольцы неосознанно оценивали мимическую экспрессию предъявляемого лица. Напомним, что в фазе 1 предъявляли лица, ориентированные прямо, а взгляд «глаза в глаза», как известно, играет важную роль в социальном взаимодействии людей и может вызвать активацию в данной области мозга [Senju et al., 2009].

В задаче определения мимики (фаза 3—фаза 4) мы наблюдаем совершенно отличную картину (рис. 2). С одной стороны экспериментальная парадигма остается такой же: меняется структура изображения, но не инструкция. Соответственно, можно было бы предположить, что мы увидим такие же закономерности в картах активациях, как и при выполнении задачи распознавания поворота лица. Но этого не происходит. Более того, мы наблюдаем существенные изменения во всей сети распознавания лица. Так, мы наблюдаем активацию в теменной коре (BA7, BA40, BA31) только при обратном вычитании, когда из фазы 4 (предъявление лиц с выраженной мимической экспрессией и поворотами) вычитаем фазу 3 (предъявление лиц с нейтральной эмоцией и ориентированных прямо).

Для того чтобы разобраться, почему в задаче определения эмоции мы видим столь существенную разницу в сравнении с выпол-

нением задачи распознавания поворота мы провели отдельную психофизическую серию исследования по выявлению порогов распознавания мимики и поворота оптоклонов. Использовали шкалу выраженности признака в программе FaceGen (Singular Inversions, Canada) от 0 (отсутствие признака) до 10 (максимально выраженный признак). Кривые радости и грусти были уравнены между собой по количеству измененных пикселов (по параметру RGB), рассчитанных для каждой градации эмоции. Подобная процедура была проведена и для поворота направо и налево. Результаты исследования на группе добровольцев представлены на рис. 5.

В целом, оценивая динамику хода кривых, можно отметить схожий паттерн. Чем сильнее выражен признак, тем более определённым по содержанию является стимул, и тем легче респонденту опознать его. И, наоборот, чем слабее выражен признак, тем вероятность правильного распознания объектов уменьшается, достигая порога случайного угадывания в условиях предъявления лиц с нейтральной экспрессией, прямо ориентированных.

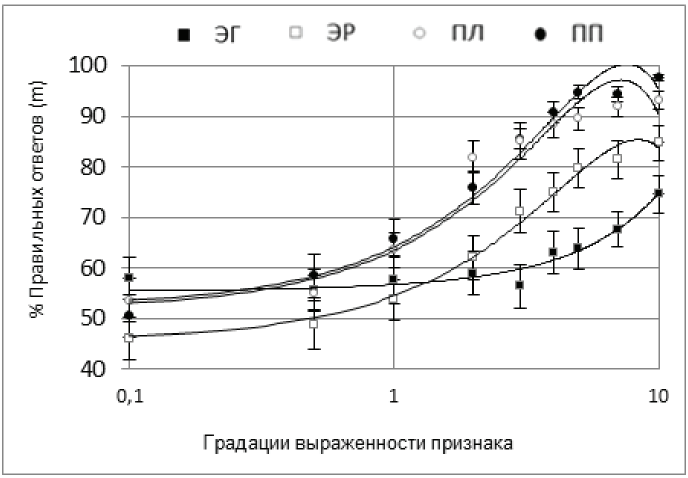


Рис. 5. Количество правильных ответов испытуемых при распознавании эмоции радости (ЭР), эмоции грусти (ЭГ), поворота направо (ПП), поворота налево (ПЛ) в зависимости от градации выраженности признака по шкале FaceGen. Данные приведены с указанием ошибки среднего значения. На графиках представлены полиномиальные линии тренда

В тоже время, видно, что среднее значение правильных ответов в задачах определения поворота выше показателей распознавания мимики лица. Более того, парный критерий Стьюдента paired t- test для сравнения средних величин двух связанных выборок показал, что в отличие от поворотов происходит статистически значимо расхождение кривых радости и грусти на уровне 4 и 5 градации выраженности признака. Порог распознавания мимических экспрессий соответствует градации 2 для эмоции радости (58,6 %) и практически градации 10 для эмоции грусти (60 %). Таким образом, несмотря на выравненность структуры изображения по количеству измененных пикселей, добровольцы эмоцию грусти распознают хуже, чем эмоцию радости.

Еще более сложная ситуация предстоит в оценивании нейтрального лица — «начала координат любых экспрессий» [Барабанщиков, Хозе, 2012]. Необходимо отметить, что изучение слабых экспрессий связано с именем известного американского и немецкого психолога Эгона Брунсвика [Brunswik, 1956]. Исследователь использовал 189 схематических изображений нейтральных лиц, составленных только из линий и эллипсов. Каждое лицо оценивалось с помощью полярных шкал (например, радостный — грустный, умный — глупый). Результаты показали, что зависимость представлений о возрасте, красоте, интеллекте зависят от расположения условного рта, расстояния между глазами и длиной носа. Так, например, наиболее радикальные изменения впечатления от восприятия нейтрального лица наблюдается в случае расположения рта: чем выше расположен рот, тем радостнее и моложе выглядит лицо, но ниже кажущийся интеллект. Сходный эффект дают широко расставленные глаза и короткий нос. Очень длинный нос во всех случаях вызывает отрицательное отношение к лицу, а высокий лоб — положительное. Это означает, что конфигуративные связи схематического лица непосредственно включены в порождение впечатлений и об экспрессиях, и об индивидуально-психологических особенностях их носителя [Барабанщиков, Хозе, 2012].

Данные исследования подтверждаются результатами распознавания нейтральных лиц программой FaceReader. Согласно представленным данным мимика нейтрального лица также обладает признаками, характерными для других эмоций. Таким образом, можно

предположить, что доброволец, оценивая стимулы в двух фазах стимуляции, оценивал их как эмоциональные изображения, что могло вызвать активацию одного и того же паттерна крупномасштабной нейронной сети. Согласно методике проведения фМРТ, карты активации рассчитываются на основе вычитания одной фазы из другой. Соответственно, на рисунках области, активируемые одинаково в обоих режимах, не видны и поэтому не окрашены. Очевидно, что это является фундаментальным ограничением метода.

Единственные области, которые были активированы при вычитании фаза 4 — фаза 3 — зоны теменной коры (BA40, BA7). Теменная кора играет ключевую роль в пространственном внимании, описании пространства, пространственных отношениях, отношениях между индивидуумами, индивидуумом и добычей [Hyvarinen, Shelepin, 1979]. Можно предположить что доброволец, выполняя инструкцию определения мимики, неосознанно оценивал также и ориентацию лица в пространстве.

Вторая рассматриваемая экспериментальная парадигма — это сравнение двух условий (фаз стимуляции), при которых тип изображения оставался одним и тем же, но менялась инструкция: в обеих фазах стимуляции предъявляли либо только нейтральные лица, ориентированные прямо, либо только лица с мимической экспрессией, повернутые в разные стороны. Таким образом, менялась инструкция, но не структура изображения. Необходимо отметить, что при вычитании фаза 2 — фаза 4 мы имеем дело со стимулами, в которых присутствует оба изучаемых признака: и эмоция и поворот. А, при вычитании фаза 1 — фаза 3 мы имеем дело с изображением, где оба признака отсутствуют. Полученные в процессе вычитания карты активации представлены на рис. 6.

В условиях предъявления лиц с мимической экспрессией, ориентированных в разные стороны (фаза 2 — фаза 4), мы наблюдаем два паттерна активации, обозначенные на рисунке разными цветами. Желтым цветом обозначена активация при выполнении задачи поворота, а голубым — при выполнении задачи эмоции. Напомним, что в стимулах, предъявляемые в обеих фазах стимуляции присутствовали оба явно выраженных признака: и поворот и эмоция. Можно предположить, что при наличии обоих очевидных признаков, проис-

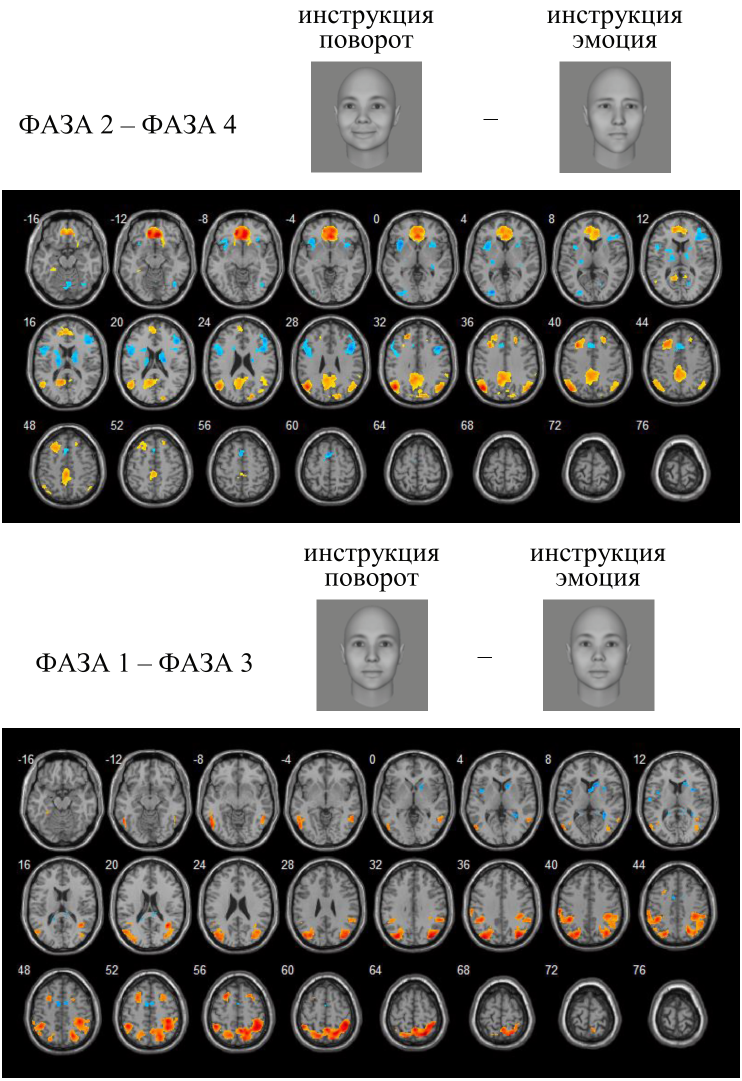


Рис. 6 — Карты активации, полученные при вычитании двух фаз стимуляции: фаза 2 — фаза 4, в обеих фазах предъявляли эмоциональные лица, повернутые в разные стороны; фаза 1 — фаза 3, в обеих фазах предъявляли нейтральные лица, ориентированные прямо. Менялась только инструкция. Данные усреднены на группе добровольцев, N=24 (без FDR-corr., $p < 0.001$)

ходит автоматическое их сравнение и ранжирование по каким-либо признакам (степени релевантности, значимости, сложности). Видимо, с этим связана активация паттерна сети базового режима в задаче определения поворота. В задаче же определения эмоции мы наблюдаем активацию в областях мозга, отвечающих за эмоциональное восприятие стимулов (например, в островковой доле конечного мозга — инсуле). Роль инсулы в эмоциональном восприятии, особенно при работе со стимулами, вызывающими негативные эмоциональные реакции, подтверждена в многочисленных исследованиях [Quarto et al., 2016]. По данным субъективных отчетов испытуемых оптоклоны воспринимались без удовольствия. Подобную негативную эмоциональную окраску чаще всего вызывал тот факт, что лица не похожи на реальных людей. Также наблюдается активация в путамене — области мозга, которая, предположительно связана с обучением с подкреплением, подготовкой и управлением двигательной активностью, участием в когнитивных функциях. Вероятно, активация данной области в задаче определения эмоции связана с тем, что данная задача была более сложной для добровольцев (рис. 3–4).

В условиях предъявления лиц с нейтральной эмоцией, ориентированных прямо (фаза 1 — фаза 3), оба изучаемых признака отсутствовали: не было ни эмоции, ни поворота. По сути, в данной фазе был применен некий аналог проективного теста, основанный на необходимости оценивать неопределенную и многозначную ситуацию. В итоге мы наблюдаем два паттерна активации. Оранжевым цветом — отмечена активация в задаче распознавания поворота, что, ожидаемо, вызвало активацию дорсальных зрительных путей, ответственных за описание пространственных отношений. А голубым цветом — отмечена активация в задаче распознавания эмоции, что повлекло активацию областей, ответственных за эмоциональное восприятие стимулов, характерных для наших условий стимуляций (активация островковой доли конечного мозга).

Заключение. Таким образом, на данном этапе удалось установить, что в различных условиях стимуляции, например в условии, когда менялась структура изображения, но не инструкция и наоборот, когда менялась инструкция, но не структура изображения, мы наблюдаем активацию различных паттернов крупномасштабной

нейронной сети головного мозга человека. Перестройка нейронной сети определяется самой задачей, стимулами–изображениями, их физическими, геометрическими и семантическими характеристиками, и, наконец, структурами глаза и зрительного мозга, обеспечивающими восприятие, узнавание и понимание и, естественно, планирования ответных действий. При наличии нескольких условий стимуляции (например, при наличии и поворота и эмоции) вероятно, происходит их автоматическое сравнение и оценка их физических и семантических характеристик. Причем данное сравнение определяется не только инструкцией, но и другими параметрами. Так, в задаче распознавания эмоции мы наблюдаем активацию также и дорсальных зрительных путей, ответственных за описание пространственных отношений, а в задаче распознавания поворота — наблюдаем активацию в медиальной префронтальной коре, ответственной, в том числе и за осознанное и неосознанное сопереживание текущему эмоциональному состоянию другого человека.

Список литературы:

Барабанщиков В. А., Хозе Е. Г. Конфигуративные признаки экспрессий спокойного лица // Экспериментальная психология. 2012. Т. 5. № 1. С. 45–68

Жукова О. В., Шелепин Ю. Е., Васильев П. П. Крупномасштабные нейронные сети восприятия мимики лица // Когнитивная психология: методология и практика. СПб.: ВВМ. 2015. С. 17–23

Шелепин Ю. Е., Борачук (Жукова) О.В., Пронин С. В., Фокин В. А., Хараузов А. К., Логунова Е. В., Васильев П. П. Лицо человека и нейрофизиология невербальных средств коммуникации. //Коллективная монография «Лицо человека в науке, искусстве и практике». Под ред. Ананьевой КИ, Барабанщикова В. А., Демидова А. А.. Москва, 2015

Шелепин Ю. Е., Жукова О. В. Нейрофизиологические исследования паттернов невербальной коммуникации // Психофизиологические и нейролингвистические аспекты процесса распознавания вербальных и невербальных паттернов коммуникации. СПб., Изд.: ВВМ. 2016. С. 17–55

Шелепин Ю.Е., Фокин В.А., Меньшикова С.В., Борачук О.В., Коскин С.А., Соколов А.В., Пронин С.В., Хараузов А.К., Васильев П.П., Вахрамеева О.А. Методы иконики и методы картирования мозга в оценке функционального состояния зрительной системы // Сенсорные системы, том 28, № 2, стр.63–78, 2014.

Brunswik E. Perception and representative design of psychological experiments. Berkley: University of California Press. 1956.

Bryan T.D., Hedy K., Kevin N.O. A Meta-Analysis of Functional Neuroimaging Studies of Self and Other Judgments Reveals a Spatial Gradient for Mentalizing in Medial Prefrontal Cortex // J. Cogn.Neurosci. 2012. V.24 (8). pp. 1742–1752.

Hyvarinen J., ShelepinYu.E. Distribution of visual and somatic functions in the parietal associative area 7 of the monkey // Brain Research. 1979. V. 169. P. 561–564.

Keuken M.C., Müller-Axt C., Langner R., Eickhoff S.B, Forstmann B.U., Neumann J. // Brain networks of perceptual decision-making: an fMRI ALE meta-analysis // Frontiers in Human Neuroscience, June 2014, Volume 8, Article 445, pp.: 1–14.

Quarto T., Blasi G., Maddalena C., Viscanti G., Lanciano T., Soleti E., Mangiulli I, Taurisano P., Fazio L., Bertolino A., Curci A. Association between Ability Emotional Intelligence and Left Insula during Social Judgment of Facial Emotions // PLOS ONE | DOI:10.1371/journal.pone.0148621

Raichle, M.E., MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. // A default mode of brain function. // Proc. Natl. Acad. Sci. U.S.A. 2001. 98: 676–682

Rizzolatti, G., and Craighero, L. The mirror-neuron system // Annu. Rev. Neurosci. V. 27. pp. 169–192. 2004.

Senju A., Johnson M.H. Eye contact effect: mechanism and development // Trends Cogn Sci. 2009. V. 13 (3). P. 127–134.

Глава 9.
Активность нейронных сетей головного мозга человека до, во время и после инсайта при распознавании изображений

К. Ю. Шелепин, В. А. Фокин, П. П. Васильев, А. В. Соколов, Г. Е. Труфанов

Технологические возможности позволили начать экспериментальное изучение деятельности нейронных сетей как при осознаваемом, последовательном, аналитическом методе решении задач, так и при внезапном скачкообразном «озарении» или «инсайте». Цель данной работы изучить перестройку нейронных сетей при переходе от неосознаваемых процессов обработки информации в осознаваемые процессы во время «инсайта». Благодаря применению стандартизованного программного синтеза динамических изображений (тестов) и цифрового анализа изображений откликов мозга установлена закономерность перестройки активности крупномасштабных нейронных сетей головного мозга человека в момент «инсайта». Активность крупномасштабных нейронных сетей максимальна в момент «инсайта», соответствующего порогу распознавания, решения задачи, и меньше при подпороговых и даже надпороговых условиях наблюдения динамических изображений. Установлено максимальное значение BOLD (Blood-Oxygen-Level Dependent) сигнала при достижении порога в одних зонах и максимальное снижение активности нейронных сетей в других зонах мозга.

Развитие нейротехнологий в последние годы, создание интерфейсов человек-компьютер, искусственных автономных систем, обеспечивающих целенаправленную деятельность и принятие решений, потребовало изучения нейрофизиологических исследований инсайта. Особый интерес представляет моделирование инсайта и деятельности нейронных сетей в задачах, решаемых не аналитическим способом для создания систем искусственного интеллекта. Оказалось, что некоторые инженерные решения созвучны тради-

ционным представлениям гештальт-психологии. Именно в рамках гештальт-психологии Дэвид Филд и Роберт Гесс предложили (рис.1) новые методы исследования инсайта, легко моделируемые нейронной сетью, имитирующей работу зоны BA17 (Brodmann area 17) первичной зрительной коры [Hess., Field.,1999; Field., Hayes., Hess F., 1993].

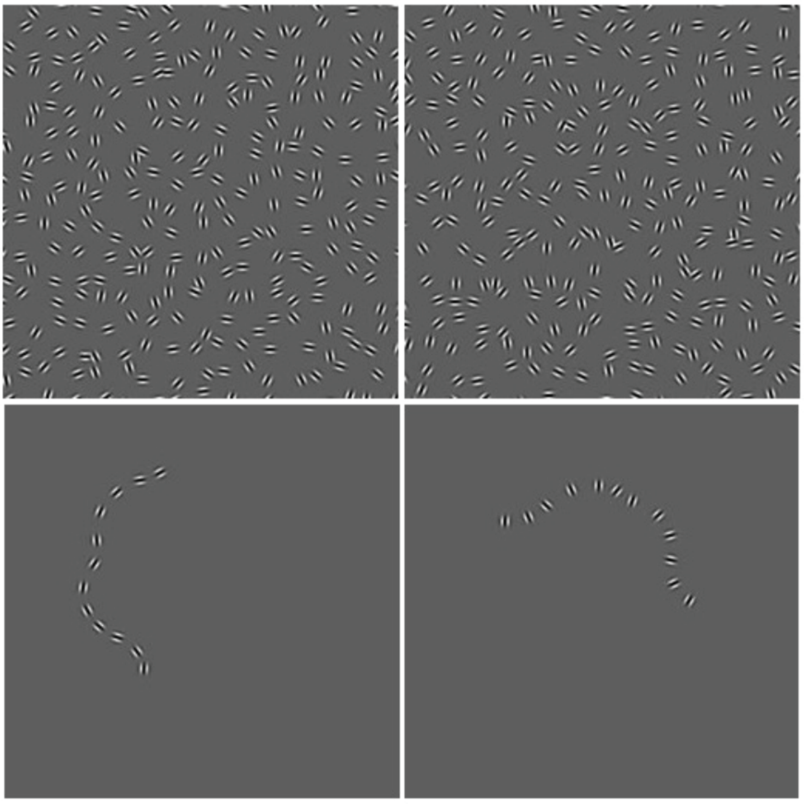


Рис. 1. Мысленное формирование «ленты» из элементов Габора, сцепленных одной «судьбой» (по терминологии гештальт психологии), происходит легко (левый верхний рисунок), так как элементы Габора, образующие ленту, имеют почти коллинеарную ориентацию. В случае неколлинеарных элементов Габора (верхний правый рисунок) восприятие гештальта — ленты, как целостной фигуры, без обучения затруднено. Подсказка — два нижних рисунка [Hess, Field, 1999]

Работы Гесса и Филда позволили отождествить порог выделения целостной фигуры в условиях неопределенности как модель зрительного инсайта. Но на практике изучения возникновения инсайта в условиях неопределенности необходимо было предложить новый метод, позволяющий измерять порог распознавания и пригодный для фМРТ. Следует заметить, что не всякий психофизический порог является аналогом инсайта. Эффект неопределённости, этап предварительного накопления и эмоциональное сопровождение порога распознавания являются важными атрибутами инсайта. Эти факторы отсутствуют при определении многих психофизических порогов, например, при определении порогов контрастной чувствительности или порогов разрешающей способности.

В качестве инструмента изучения инсайта нами был выбран метод определения порога распознавания неполных фрагментированных изображений, компьютеризированный метод Голлин-теста [Foreman, Hemmings, 1987] в версии, разработанной в лаборатории физиологии зрения ИФ РАН [Шелепин и др., 2015]. Схема предъявления представлена на рис. 2. Важно отметить то, что для исследования изменения во времени BOLD сигнала применили изменяющийся во времени тестовый зрительный сигнал. Существующие методы визуализации активности головного мозга человека и разработанный нами метод цифрового синтеза тестовых сигналов (неполных изображений, соответствующих Голлин-тесту) позволили определить пути решения поставленной задачи [Шелепин и др., 2015; Шелепин и др., 2017]. Данный стимул для распознавания требует вовлечения последующих, менее изученных отделов мозга [Глезер, 1995].

Таким образом, достижение цели на данном этапе сведено к решению задачи определения порога, соответствующего моменту перехода результатов неосознаваемых, подсознательных процессов на осознаваемый уровень принятия решений. Этот переход может быть отождествлен с моментом возникновения инсайта, общеизвестного явления внезапного понимания решаемого процесса, в данном случае сведения разрозненных фрагментов в единое целое, в гештальт путём «внезапного озарения».

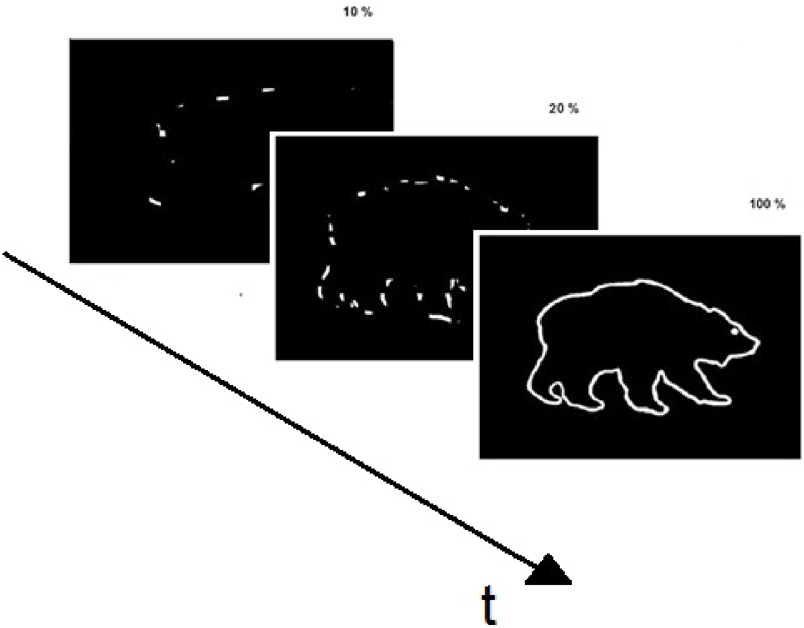


Рис. 2. Отдельные кадры из последовательного ряда тестовых изображений, заполняющих контур изображения Голлин-теста отдельными фрагментами в течение одной минуты. На рисунке даны примеры фрагментированного 10, 20 и 100 % заполнения контура тестового изображения [Foreman, Hemmings. 1987; Шелепин и др., 2015]

Проведённая нами работа позволяет утверждать, что Голлин-тест, как один из наиболее эффективных и изученных методов тестирования мозга, является удобным и надёжным методом для изучения механизмов возникновения инсайта. Ранее нами было показано, что порог распознавания в среднем при первом предъявлении равен 20 %, но меняется и зависит от сложности рисунка, знакомства испытуемого с данным рисунком и с методом исследования, от состояния испытуемого и его возраста. Для применения Голлин-теста, как метода, имитирующего инсайт, важно то, что момент распознавания сопровождается выраженной осознаваемой эмоциональной реакцией, регистрируемой нами отдельно [Foreman, Hemmings, 1987; Шелепин и др., 2015; Шелепин и др.,

2017]. Процесс достижения инсайта протекает неосознанно. Наблюдатель видит фрагменты, ожидает увидеть объект, но не знает какой, поэтому работает в условиях неопределённости. Поэтому он не может и не строит аналитических, логических схем решения задачи. Достижение порога происходит неожиданно и вызывает естественную эмоциональную реакцию разной степени выраженности. Следовательно, можно утверждать, что разработанный нами инструмент позволяет исследовать нейрофизиологические механизмы инсайта.

Цель данной работы – изучить изменения BOLD-сигнала по данным функциональной магнитно-резонансной томографии (фМРТ) суммарной активности головного мозга человека и зон, входящих в крупномасштабные нейронные сети, обеспечивающие порог распознавания и перераспределение активности этих сетей до, при и после инсайта, отождествляемого с достижением порога распознавания неполных фрагментированных изображений тестовых объектов.

В качестве динамических изображений выбран компьютеризированный Голлин-тест, запуск которого приводит к постепенно нарастающему контуру из случайно возникающих фрагментов [Foreman, Hemmings, 1987; Шелепин и др., 2015; Шелепин и др 2017].

Для измерения состояния мозга в момент возникновения инсайта мы использовали фМРТ, которая позволяет определять активацию разных областей мозга на основании изменения локального кровотока. В качестве стимулов был использован динамический сигнал — постепенное во времени заполнение контура различных контурных изображений живых и неживых объектов. Выбор данного теста, как модели вызова инсайта, был обусловлен возможностью контролировать подпороговые, пороговые и надпороговые состояния мозга в процессе восприятия нарастания процента заполнения контура. Регистрация психоэмоциональной реакции в период до возникновения порога, в момент порога и после осуществлялась при помощи самотестирования испытуемого, электроэнцефалографии (ЭЭГ) и электромиографии (ЭМГ) лицевых мышц. Также регистрировались изменения диаметра зрачка и микротремор глаз.

Предъявление нарастающего числа элементов контура

При выполнении распознавания Голлин-теста исследовали карты распределения мозговой активности по BOLD-сигналу в момент предъявления стимула – динамического изображения относительно распределённой активности нейронной сети базового режима мозга (относительного покоя – REST). Сравнение осуществляли на основании стандартных методов статистической обработки изображений в среде MatLab и SPM12 «виртуальных срезов мозга» по данным фМРТ.

Мы сравнивали изображения, отражающие активность разных зон мозга, не только относительно «покоя», но и как разницу активности в момент достижения порога распознавания относительно допорогового состояния и в момент достижения порога распознавания относительно послепорогового состояния мозга. Особенность фМРТ-измерений состоит в том, что интерес представляет не абсолютное значение BOLD-сигнала, а разница между его значением в разных состояниях испытуемого. Таким образом, в этих условиях мы всегда имеем дело с относительной величиной, хотя и можем оценивать активность той или иной структуры в числе активированных вокселов, аналогов пикселов для трехмерного пространства. Получены статистически достоверные различия максимальной активности мозга в момент достижения порога, в момент, когда наблюдатель внезапно видит в случайном паттерне точек или фрагментов контура целостный объект, что и является сутью инсайта. Сравнение карт разниц состояния активности при пороге от до порогового и после порогового состояния, представленных на рис. 3А, 3Б, 3В, демонстрирует, что различия видимы даже визуально, «на глаз». На графиках, это видимое на картах различие активности мозга в разные моменты, представлены количественно на рис. 4. На гистограммах (рис.4) в центре виден значимый пик. Этот пик на графиках соответствует моменту порогового значения распознавания формы неполного, фрагментированного изображения. Последующее увеличение количества проявляемых на экране фрагментов контура тестового изображения уже не увеличивает различие с «покоем», а напротив, уменьшает это различие. На рис. 3А, 3Б и 3В показан результат статистической обработки разницы состояния активности мозга до, во время и после возникновения инсайта относительно состояния «покоя» по данным фМРТ.

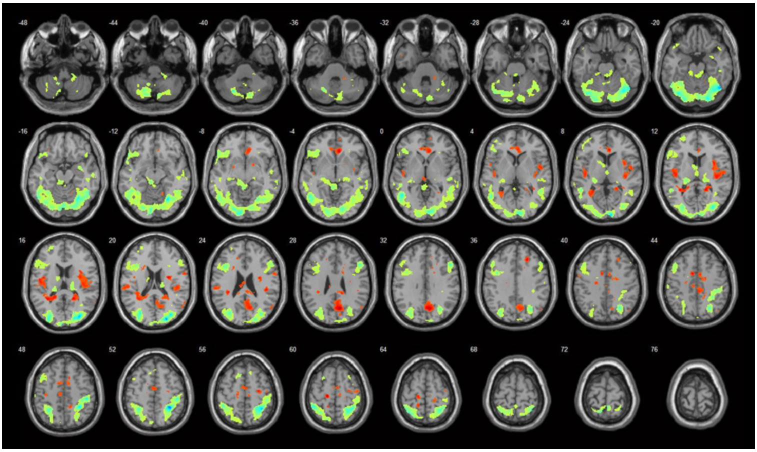


Рис. 3А. Усреднённый уровень активации для группы из 22 человек (FDR-corrected, $p < 0{,}01$) в задачах на распознавание неполно представленного изображения объекта. Зоны мозга, окрашенные в красный цвет, – результат статистической оценки активности мозга в состоянии «покоя» относительно состояния «до инсайта» при предъявлении 10 % контура. Зоны мозга, окрашенные в зелёный цвет, – состояние мозга «до инсайта» при предъявлении 10 % контура относительно состояния «покоя»

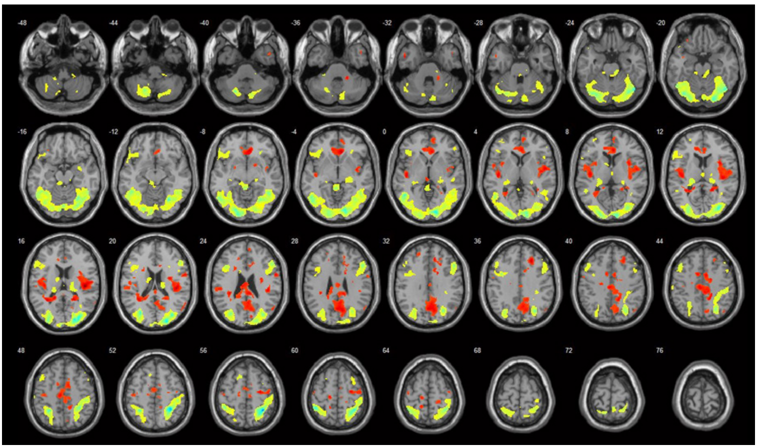


Рис. 3Б. Усреднённый уровень активации для группы из 22 человек (FDR-corrected, $p < 0{,}01$) в задачах на распознавание неполно представленного изображения объекта

Зоны мозга, окрашенные в красный цвет, – результат статистической оценки активности мозга в состоянии «покоя» относительно состояния «инсайта» при предъявлении 20 % контура. Зоны мозга, окрашенные в зелёный цвет, – состояние мозга «инсайта» при предъявлении 20 % контура относительно состояния «покоя».

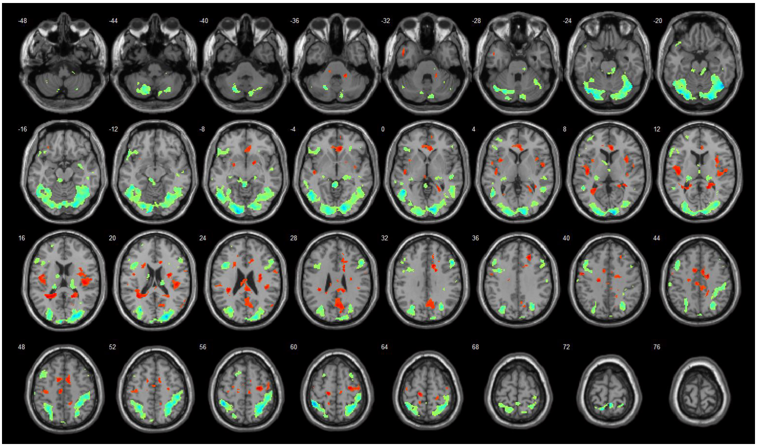


Рис. 3В. Усреднённый уровень активации для группы из 22 человек (FDR-corrected, $p < 0{,}01$) в задачах на распознавание неполно представленного изображения объекта.

Зоны мозга, окрашенные в красный цвет, – результат статистической оценки активности мозга в состоянии «покоя» относительно состояния «после инсайта» при предъявлении 60 % контура. Зоны мозга, окрашенные в зелёный цвет, – состояние мозга «после инсайта» при предъявлении 60 % контура относительно состояния «покоя».

На рис. 3А и 3В в условиях до, при и после достижения порога распознавания видна активность слева в лобной доле практически на всех срезах. Но на рис. 3Б видно, что увеличивается число вокселов в правом полушарии в лобной доле срезов 20–30 и в затылочной доле срезов от 16 до 0 и срезов от 16 до 24. Это важный результат, он согласуется с общими предположениями о роли правого полушария в развитии инсайта.

На полученных нами с помощью фМРТ «срезах мозга» в момент инсайта активация зон мозга на основании статистического сравне-

ния больше по сравнению с покоем и со значением активации при подпороговой и надпороговой стимуляции.

На рис. 4 представленны усредненные данные по всему мозгу.

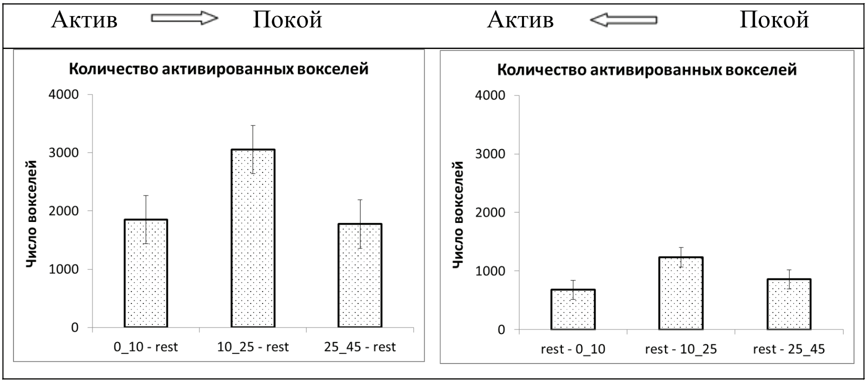


Рис. 4. Средние значения разницы активированных вокселов «зон интереса» головного мозга при двух типах статистического сравнения: «активация» относительно «покоя» (слева) и «покоя» относительно «активации» (справа). Столбики, с указанием стандартной ошибки среднего значения (SE), отражают эту разницу до появления инсайта, во время инсайта и после достижения инсайта. Состояние мозга (по данным фМРТ) до начала стимуляции условно называют «покой или rest» [Шелепин и др., 2017]

Важно отметить, что в момент достижения порога распознавания, который мы отождествляем со своеобразным инсайтом, наблюдается максимум общего числа активированных вокселов фМРТ-срезов. Это связано с тем, что наблюдаются локальные особенности уровня активации, которые значительно ниже в одних зонах и выше в других цитоархитектонических зонах, входящих в разные нейронные сети. Наглядно это увидим при наблюдении динамики изменения BOLD-сигнала во времени.

Изучили активность крупномасштабных нейронных сетей, связывающих структуры префронтальной коры (BA9, BA10, BA11), височно-теменной коры (BA39, BA40, BA22), лимбической коры (BA23, BA24, BA29, BA30, BA31), задневисочной (BA37) и затылочной коры (BA17, BA18, BA19) и подкорковых ядер (инсулы, амигдалы и др.). Показали, что существуют некоторые различия в акти-

вации левого и правого полушария. Первые 15 с после начала стимуляции они одинаковы и слева и справа, если рассматривать суммарно ответ по BOLD-сигналу всех областей. Хотя и имеются отличия реакции определённых зон на первые, как кажется наблюдателю, случайно появляющиеся точки — фрагменты будущего контура. Затем начинается отклик на формирование из случайных скоплений точек некоего паттерна. Этот отклик происходит не во всех областях. В одних он проявляется увеличением активности BA37, BA19, а в других снижением активности, как например, в медиальной теменной области BA7 или Precunius. Все поле 7, включая Precunius, очень разнородно, его разные зоны включены в разные нейронные сети, и каждая из них выполняет разные функции.

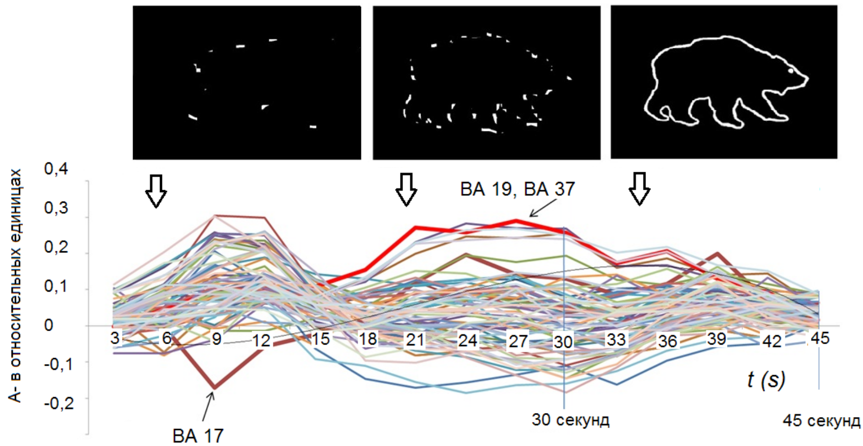


Рис. 5. Изменение во времени BOLD-сигнала большинства зон мозга по Бродману. По оси абсцисс – время и одновременно процент заполнения контура, численно равный моменту предъявления. По оси ординат: A — величина BOLD-сигнала в относительных единицах, далее называемая как Z — оценка. Стрелка указывает среднее время и численно соответствующий процент заполнения контура

На рис. 5 видно изменение BOLD-сигнала практически во всех исследуемых зонах мозга во времени в соответствии с изменением входного сигнала. Входной сигнал до пороговый, пороговый и после пороговый условно изображен в виде трех отдельных кадров

вверху рисунка. Стрелки указывают какому моменту времени на шкале соответствует демонстрируемое тестовое изображение на стрелками. Но после уже 35 % предъявления контура большинство изображений воспринимается как узнанное изображение и активность BOLD-сигнала затухает. Из рисунка видно, что в момент неопределенности (до пороговый) большинство зон увеличивают активность.

Второй важный вывод из этого довольно запутанного графика состоит в том, что в момент распознавания, или завершения периода неопределенности одни участки мозга повышают свою активность, а другие снижают. Это самый интересный результат, так как изменение откликов разных зон мозга на изменение во времени входного сигнала отражают его работу.

Рассмотрим отклики мозга отдельных областей головного мозга. Обычно инерционность BOLD- сигнала не позволяет точно определять развитие отклика на тестовый сигнал в той или иной области головного мозга человека. Но мы предъявляли отдельные фрагменты со скоростью один фрагмент в секунду, «замедлили время» стимуляции. Этот экспериментальный прием позволил увидеть динамику изменения BOLD-сигнала.

На рис. 6–9 мы видим синхронность изменения BOLD-сигнала после начала предъявления. Сразу после начала предъявлений характерно торможение только участков первичной зрительной коры BA17 и BA18 и «безучастность» в этот момент зон BA19, BA37. Но затем происходит неуклонное нарастание BOLD-сигнала в первичных областях зрительной коры BA17, BA18 и значительный рост на последующих этапах обработки в BA19 и BA37.

Максимум ответа – в зоне BA37, и этот максимум совпадает с моментом распознавания, т. е. при возникновении инсайта. Мы видим, что справа оппонентные отклики зоны BA37 и прекунеус более выражены, чем слева. Пик расхождения совпадает с моментом возникновения инсайта или с порогом распознавания и максимумом эмоциональной реакции испытуемого.

С нашей точки зрения перистриарная кора (которая включает зоны по Бродману BA19 и BA37) является ключевой для формирования целостного образа — гештальта. Ранее знали, что эти зоны

мозга обеспечивают формирование конкретного образного мышления [Глезер, 1993]. Последствия разрушения BA37 соответствует классическим клиническим представлениям об агнозиях [Кок, 1967].

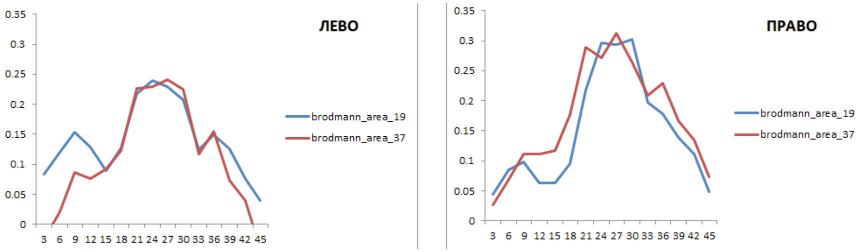


Рис. 6. Изменения BOLD-сигнала в затылочных зонах по Бродману BA19 и BA37. (левое полушарие, правое полушарие). По оси абсцисс – время и соответствующей ему процент предъявления контура. По оси ординат – значение BOLD-сигнала в относительных единицах Z оценки

Отклики работы в затылочно височной части перистриарной коры в BA37 ведут себя оппонентно к откликам в теменной коре. Это представлено на рисунке 7. Видно что в момент максимума отклика в зонах BA19 и BA37, соответствующего развитию «инсайта» и пониманию какое изображение можно мысленно «склеить» в целостный образ, в гештальт, происходит относительное торможение активности в теменной коре BA7 и ее медиальной части. Однако моменту торможения предшествует активация в теменной коре и в лобных отделах, соответсвующая широко известным данным о роли внимания в теменной коре и в лобной доле.

Рис. 8 достаточно наглядно демонстрирует сложную перестройку ответов теменных и лобных областей мозга на динамически развивающийся стимул. Интересно и возникновение активации в височной области в слуховой зоне BA22 и нижневисочных долях BA21 и BA20 связанных с обработкой сложной зрительной информации.

По оси абсцисс – время и соответствующей ему процент предъявления контура. По оси ординат – значение BOLD-сигнала в относительных единицах.

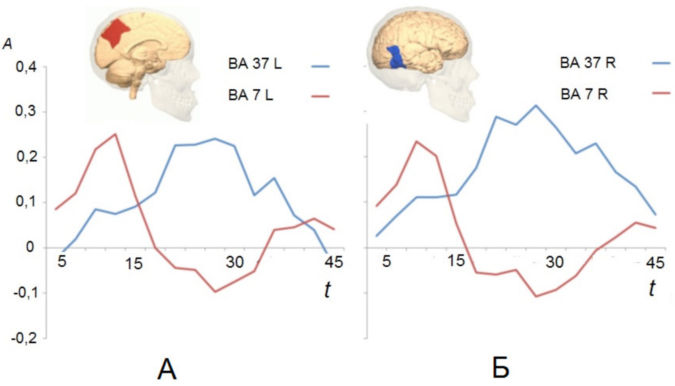


Рис. 7. Отклики: затылочно-височной коры BA37 и медиальной области теменной коры (BA7), на участке так называемого предклинья (precuneus). Максимальное различие откликов этих оппонентных зон наблюдается в момент порога распознавания. А- левое полушарие, Б — правое полушарие. По оси абсцисс – время и соответствующей ему процент предъявления контура. По оси ординат – значение BOLD-сигнала в относительных единицах Z оценки

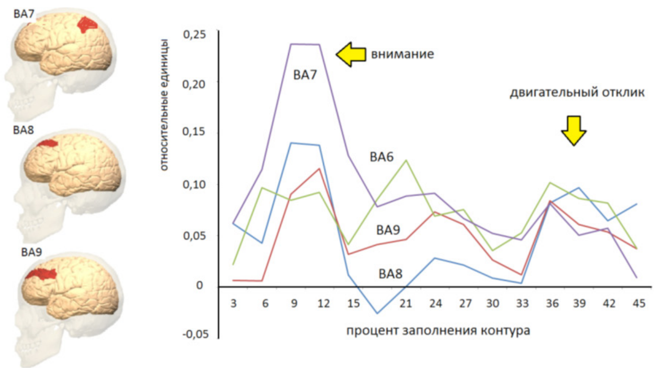


Рис. 8. Отклики теменных (латеральной части BA7) и лобных (BA8, BA9) долей мозга, связанных вначале с ориентировочной реакцией на появление стимула (вниманием) малыми изменениями при распознавании и синхронной активацией после распознавания. По оси абсцисс – процент предъявления контура и численно соответствующее ему время в секундах. По оси ординат – значение BOLD-сигнала в относительных единицах Z оценки

Рассмотрим изменения BOLD-сигнала в лобных долях. Видно, что в левом и в правом полушариях в зонах BA45 и BA46 выраженно развивается реакция на появление сигнала. Справа BA46 активируется сильнее, в момент, когда развивается «инсайт». Этот ответ возникает несколько позднее срабатывания крупномасштабной нейронной сети, обеспечивающей распознавание, и включающей BA37. Нейронные сети в лобных долях работают своеобразно. Зона BA9 и BA10 демонстрируют слабые отклики, но значимые отличия. Активность зоны BA9 не изменяется, а зона BA10 тормозится. Торможение активности зоны BA10 лобной коры в момент развития инсайта значимо сильнее в левом полушарии.

Верхний ряд графиков на рис. 9 показывает активацию зон, выбранных нами, как зоны интереса, отражающих общее внимание при предъявлении стимула. Нижний ряд графиков на рис. 9 показывает активацию выбранных цитоархитектонических зон лобных долей, в момент инсайта, участвующих в момент распознавания изображения и принятия решения.

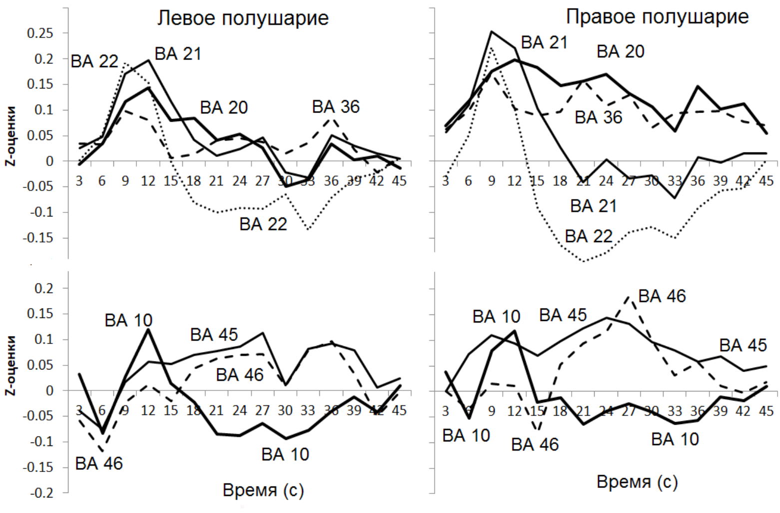


Рис. 9. Примеры откликов височных зон (BA 20, 21, 22, 36), активированных самим появлением стимула (верхние графики), и лобных зон, как проявляющих активность (BA45 и 46), так и не проявляющих ее (BA10) в момент распознавания (нижние графики)

Принятие решения в результате инсайта предполагает кажущееся внезапным, скачкообразное решение задачи задачи [Валуева, Ушаков, 2015, 2017; Ушаков, 2011]. Он отражает образное мышление, механизм которого более древний и вероятно основной, присущ и хищникам и приматам. Противоположный ему механизм — это детальный последовательный аналитический способ решения задачи. У человека этот тип мышления развит в полной мере и обеспечивает осознанное мышление на всех этапах. Инсайт — это высшее и наиболее наглядное проявление целостного, образного мышления, которое на первых этапах обработки информации проходит неосознанно.

Непосредственный анализ результатов требует обратить внимание на некоторые «побочные» изменения активности BOLD-сигнала, связанные с процедурой любого исследования фМРТ. На период подпорогового предъявления приходится первый известный подъём увеличения BOLD-сигнала на 4–10 секунд от начала предъявления. Этот подъём активности не связан с порогом распознавания и в нашем случае не связан с порогом «инсайта». Этот подъем активности наблюдается во многих зонах и развивается всегда в одно и то же время при любой стимуляции. Для большинства зон в этот момент характерно наличие активация зон, связанных с вниманием. Это теменная, височная и лобные зоны (BA7, 8; BA22, 21, 20; BA9, 10). В затылочных зонах (BA17, 18) наблюдается торможение. В перистриарной коре выраженных изменений в этот момент не наблюдается. Для того, чтобы дифференцировать интересующий нас процесс и наступающий сразу после начала предъявления любого стимула известный подъём активности, мы проводили отдельные серии исследований, в которых варьировалось время предъявления фрагментов и тем самым изменялось наступление времени порога распознавания.

Интересующая нас динамика активности мозга четко отслеживала динамику изменения тестового сигнала во времени. Мы изменяли скорость заполнения контура. Естественно, порог распознавания наступал в разное время от начала стимуляции. По данным изменения BOLD-сигнала мы впервые смогли рассмотреть нейрофизиологические механизмы реакций мозга на изменяющийся сигнал. Мы сравнили величину отклика в момент достижения порога относительно

покоя и относительно подпороговой стимуляции и относительно надпороговой стимуляции.

На основании полученных результатов ещё раз укажем, что в момент достижения порога распознавания, который мы отождествляем с инсайтом, наблюдается максимум общего числа активированных вокселов фМРТ. Интересно, что максимум различия достигается как при сравнении состояний «покой» минус «активация», так и «активация» минус «покой». Это связано с тем, что наблюдаются локальные особенности уровня активации, которые значительно ниже в одних зонах и выше в других цитоархитектонических зонах, входящих в оппонентные нейронные сети, принадлежащие к системе «покой» и к системе «активация», и нейронных сетей быстрого реагирования.

Важно, и это мы впервые определили, что после достижения порога распознавания в данной задаче в момент надпорогового предъявления стимула происходит спад активности по сравнению с моментом порога распознавания. Это, казалось бы, простое наблюдение характерно не для любого достижения порога, а только того, при котором ликвидируется неопределенность и процесс распознавания завершается. Мы зафиксировали, что при распознавании неполных контурных изображений Голлин-теста активность правого полушария превышает активность левого полушария.

Роберт Гесс и Дэвид Филд на основе анализа собственных результатов моделирования горизонтальных взаимодействий в первичной зрительной коре показали возможность восстановления контура за счёт внутрикорковых горизонтальных связей [Field, Hayes, 1993; Hess, Field,1999]. Однако предложенное ими объяснение инсайта только организацией связей коротких межмодульных связей BA17 вызывает возражения. Горизонтальная активность может развиваться в пределах стриарной коры BA17 короткими связями распространяющаяся по коре и имеющая торможение в направлениях ортогональных к другим ориентациям.

Стимулы, примененные в виде фигуры «змеи», оптимальны для демонстрации взаимодействия в пределах стриарной коры, так как следуют один фрагмент за другим. Если «змея» неколлинеарна, и мы ее пусть с трудом но видим, следовательно требуется другой меха-

низм объясняющий построение змеи из неколлинеарных элементов. Поэтому нужна была другая методика и мы её применили [Шелепин и др., 2015; 2017; 2018]. Для исследования неосознаваемых полностью процессов уже применяли фМРТ-исследования в нашем институте. Но применяемые методики не позволили увидеть процесс до, во время и после достижения порога. Применённый нами известный и разработанный в институте компьютеризированный Голлин-тест позволил изучить другие механизмы взаимодействия осознаваемых и неосознаваемых процессов обработки сенсорной информации. В результате проведённой работы удалось показать, что для объединения фрагментов в единое целое в гештальт и его распознавания и классификации недостаточна работа нейронной сети взаимодействующих модулей первичной зрительной коры. Этим другим механизмом может служить зона, обеспечивающая целостное восприятие в нижневисочной коре BA37. Что мы и видим в наших исследованиях.

Нами впервые показана динамика изменения BOLD-сигнала от подпорогового до надпорогового уровня. Измерение порогов распознавания методами цифровой обработки откликов головного мозга позволит перейти к целенаправленному моделированию искусственных нейронных сетей. Порог распознавания сопровождается максимумом активности всего мозга, превышающем уровень «покоя» при отсутствии стимулов, а также подпороговый и надпороговый уровень, когда объект был виден ясно и чётко. Перестройка активности нейронной сети обусловлена перераспределением активности оппонентных корковых областей мозга и подкорковых ядер, что было установлено на основании изменения BOLD-сигнала. Изменение BOLD-сигнала во времени коррелирует с изменением спектра ЭЭГ — снижением альфа- и повышением гамма-ритма. Этот результат подтверждает результаты других исследователей [Kounios, Beeman, 2014].

Максимум нейронной активности в момент порогового распознавания формы фигуры сопровождается выраженной эмоциональной реакцией, которую мы регистрировали на основании устных ответов испытуемых и результатов объективных измерений колебаний размеров диаметра зрачков, микротремора глаз и напряжением мимических мышц, сопровождающих любую эмоциональную реакцию

[Шелепин и др., 2015].

Видно, что активность в одних областях уменьшается и увеличивается в других. Эти изменения происходят в процессе действия стимула. Оппонентные взаимодействия, например, BA17 и BA37 или BA9 и BA10 в затылочных, височных и, главное, во фронтальных областях головного мозга отражают многоуровневый процесс принятия решений по конфликтующим оценкам, предсказанным в теоретических исследованиях [Saaty, 2008] и реализованных в оппонентных структурах головного мозга человека. Установлен максимум активности одних зон в момент достижения порогового уровня предъявления контура и его снижение после достижения порогового уровня. В оппонентных зонах пороговый момент совпадает с абсолютным минимумом активности. Обеспечение баланса критериев при принятии решения осуществляется путём перераспределения активности между областями мозга. Структуры мозга, обеспечивающие важнейшие когнитивные и непременно сопровождающие их эмоциональные процессы, взаимодействуют по принципу синхронности и оппонентности, т.е. осуществляют решения по множественным критериям и по конфликтующим оценкам [Ardila et al., 2010; Saaty, 2008].

На рис. 10 показаны примеры разворачивания ответов различных, но очень важных областей мозга на зрительный сигнал.

В задней части нижневисочной области и в фузиформной извилине расположена зона BA37. Она имеет значимые «возбудительные и тормозные связи» с левой нижневисочной областью (BA20, BA21), с левой префронтальной корой (BA9, BA46, BA45 и BA47), левой инсулой (BA13), левом и правым предклиньем — precuneus (захватывающая зоны BA7, BA19) и с затылочной областью (BA18). Вероятно, левая зона BA37 – это область узлового пересечения двух нейронных сетей: одной, осуществляющей зрительное распознавание, и другой, обеспечивающей описание семантических свойств речи [Kraft et al., 2006; Шелепин и др., 2014]. Распознавание Голлин теста вначале вызывает активацию областей связанных со вниманием пространственным (BA 7) и эмоциональным (BA 20, 21). Активируется вначале и слуховая кора. Человек настораживается в момент неопределенности. Затем наступает торможение в зонах пространственного внимания BA7, эмоционального BA20, 21 и в слуховой

зоне. Реакция слуховой коры (См. рисунок 10) периодами активации в условиях неопределенности (преинсайт) и торможения (инсайт) чрезвычайно интересно. На этом рисунке, справа, вверху даны изменения кровотока во времени в предклинье — precuneus в нижневисочной фузиформной коре части поля 37 т 19 по Бродману. Видно их выраженное оппонетное взаимодействие.

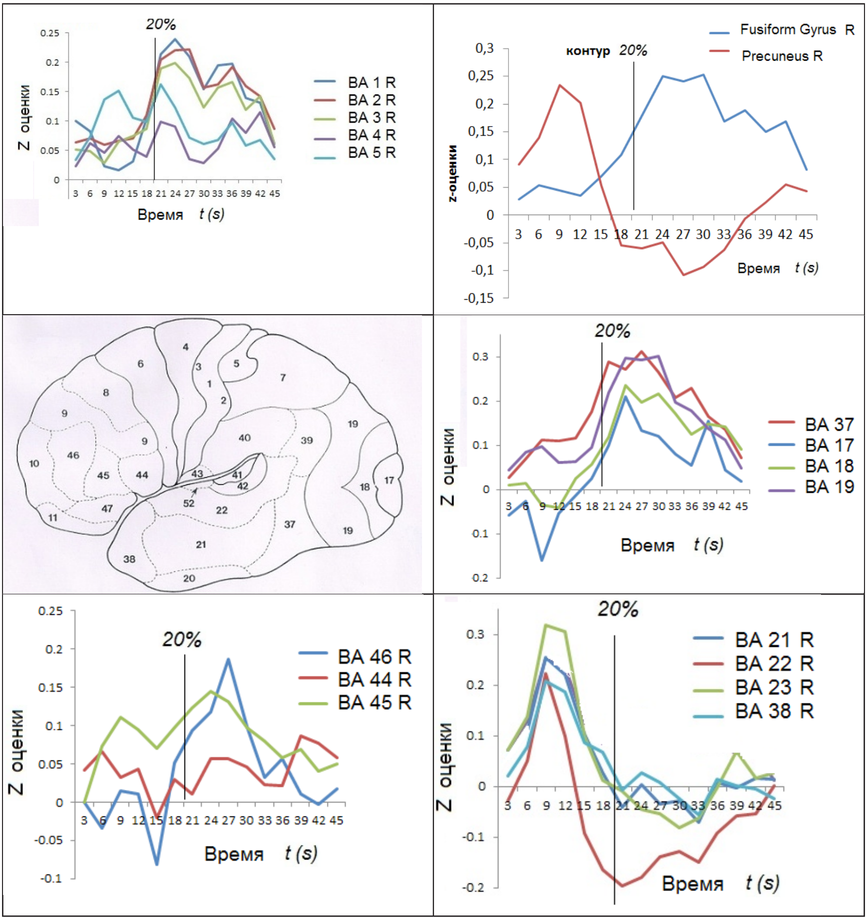


Рис. 10. Обобщающая демонстрация разворачивания во времени отклика всего мозга на динамический зрительный стимул. Показаны отклики выбранных областей лобной, теменной, височной и затылочной коры

Отношения зрительной, слуховой и соматосенсорной коры, наконец, удалось продемонстрировать как взаимодействия различных нейронных сетей, при отклике только на зрительный сигнал. Общие принципы взаимной активации и торможения при межсенсорном взаимодействии методами 30-х годов прошлого века были показаны впервые еще С. Кравковым. Соматосенсорная кора реагирует активацией после возникновения инсайта как результат получения сигнала от напряжения мышц, в первую очередь лицевых. См. рисунок 10. Учитывая инерционность метода фМРТ мы пока не может оценивать распределение активности в течении первых 500 мс после начала стимуляции. Поэтому мы и даем растянутый сигнал во времени.

Если относить инсайт к гештальту — целостному процессу восприятия, то установленные нами нейронные сети, включающие BA37, указывают конкретные нейрофизиологические механизмы, принимающие участие в обеспечении инсайта, важнейшего мыслительного явления. Его обеспечивает дорзально-фронтальное (BA19, BA37- BA45, BA46) взаимодействие, а аналитическое мышление, как обычно считают, «происходит во фронтальной коре».

Заключение и выводы

Разработанные методы позволили впервые показать отклики всего мозга на динамические зрительные сигналы. Показаны отклики на зрительные сигналы и их развитие во времени при переходе от неопределенности к распознаванию в зрительных, слуховых, двигательных соматосенсорных и в «ассоциативных» областях мозга. Нами впервые показана динамика изменения BOLD-сигнала от подпорогового, до порогового и надпорогового уровня [Шелепин, 2017, 2018]. Измерение порогов распознавания методами цифровой обработки откликов головного мозга позволит перейти к целенаправленному моделированию инсайта в искусственных нейронных сетях, осуществляющих неаналитический способ решения задач. Порог распознавания сопровождается максимумом активности всего мозга, превышающем уровень «покоя» при отсутствии стимулов, а также подпороговый и надпороговый уровень, когда объект был виден ясно и чётко. Перестройка активности нейронной сети обусловлена перераспределени-

ем активности оппонентных корковых областей мозга и подкорковых ядер, что было установлено на основании изменения BOLD-сигнала. Изменение BOLD-сигнала во времени коррелирует с изменением спектра ЭЭГ – снижением альфа- и повышением гамма-ритма. Этот результат подтверждает результаты других исследователей [Kounios, Beeman, 2014]. Максимум нейронной активности в момент порогового распознавания формы фигуры сопровождается выраженной эмоциональной реакцией, которую мы регистрировали на основании устных ответов испытуемых и результатов объективных измерений колебаний размеров диаметра зрачков, микротремора глаз и напряжения мимических мышц, сопровождающих любую эмоциональную реакцию. В задней части нижневисочной области и в фузиформной извилине расположена зона BA37. Она имеет значимые вероятно тормозные связи с левой нижневисочной областью (BA20, BA21), и активирующие с префронтальной корой (BA9, BA46, BA45 и BA47), левой инсулой (BA13), левым и правым предклиньем (захватывающим зоны BA7, BA19), и в затылочной области (BA18). Вероятно, левая зона BA37 – это область узлового пересечения двух нейронных сетей: одной, осуществляющей зрительное распознавание, и другой, обеспечивающей описание семантических свойств речи.

Результаты обработки изображений карт активности кровотока в мозге по BOLD-сигналу позволило установить, что вне затылочной (зоны BA17, BA18, BA19) и височной коры (зоны BA37) важное участие в обеспечении решения задач данного типа принимает активность инсулы (BA13), префронтальной коры (BA9, BA46, BA47, BA10) и поясной коры (BA24, BA30).

Показанный экспериментально в данной работе многоуровневый процесс принятия решения по конфликтующим оценкам, как мы видим, реализован в оппонентных структурах головного мозга человека [Шелепин и др., 2014]. Важность принятия решения по конфликтующим (оппонентным в нашей терминологии) оценкам был предсказан в известных теоретических исследованиях Саати [Saaty, 2008]. Установлен максимум активности определённых зон в момент достижения порогового уровня предъявления контура, соответствующего моменту возникновения инсайта. Это повышение активности вызывает снижение активности в других, оппонентных

зонах. Обеспечение баланса критериев при принятии решения осуществляется путём перераспределения активности между областями мозга. Структуры мозга, обеспечивающие важнейшие когнитивные функции и непременно сопровождающие их эмоциональные процессы, взаимодействуют синхронно, но между ними существуют оппонентные взаимодействия. В результате становится возможным осуществлять принятие решения по множественным критериям и по конфликтующим между собой оценкам события. Именно такой принцип взаимодействия элементов в процессе принятия решений был предсказан Саати [Saaty, 2008]. Эти самые общие представления о принятии решений в настоящее время могут быть применены для оптимизации искусственной нейронной сети как локальной, так и крупномасштабной, обеспечивающей целенаправленное мотивированное поведение. Эти представления перекликаются с рядом существующих работ по локальным [Wechsler, 2014] и глобальным нейронным сетям [Шелепин и др., 2014; Weng et al., 2013]. Моделирование деятельности нейронных сетей позволяет предполагать, что загадочный механизм инсайта может быть организован на более простой нейронной сети, присущей как высшим позвоночным хищникам, так и приматам. Загадочность этого механизма в том, что его предварительная часть происходит вне нашего сознания, и в сознание приходит уже результат предварительной обработки изображений. Аналитический способ решения задач, вероятно, гораздо более сложный, так как требует вовлечение элементов внутренней речи и логического осознанного мышления, тесно связанного с речью человека.

Обучение нейронных сетей основано именно на оппонентных взаимодействиях. Эти взаимодействия многоуровневые. Первый уровень – оппонентность внутри рецептивного поля, следующий — оппонентность между рецептивными полями и высший уровень – оппонентность между крупномасштабными нейронными сетями. В данной работе мы рассмотрели некоторые особенности именно крупномасштабных нейронных сетей. И этот тип оппонентности различается медиаторами. Обучение нейронных сетей может происходить за счёт наказания или наград, либо в реальной жизни – при своевременном или при запоздалом распознавании врага от друга.

Литература

Глезер В. Д. Зрение и мышление. Л.: Наука, 1993. 250 с.

Валуева Е. А., Ушаков Д. В. Сигнальная модель Инсайта: от исторических предпосылок к эмпирическим предсказаниям // Современные исследования интеллекта и творчества. Серия: «Экспериментальные исследования», Отв. ред. А. Л. Журавлев, В. Д. Ушаков, М. А. Холодная — 2015. Москва: Ин-т психологии РАН. С. 15–47.

Валуева Е. А., Ушаков Д. В. Инсайт и инкубация в мышлении: роль процессов осознавания / Сибирский психологический журнал. 2017. № 63. С. 19–35.

Кок Е. П. Зрительные агнозии. Л. Медицина, 1967. С.45–49.

Ушаков Д. В. Психология интеллекта и одаренности// Российская академия наук. Изд-во «Институт психологии РАН», Москва, 2011

Шелепин К. Ю., Пронин С. В., Шелепин Ю. Е. Распознавание фрагментированных изображений и возникновение “инсайта” // Оптический журнал. 2015. Т. 82. № 10. С. 70–78.

Шелепин К. Ю., Шелепин Ю. Е. Нейрофизиология «инсайта» // Петербургский психологический журнал. 2015. № 11. С. 19–38.

Шелепин К. Ю., Соколов А. В., Фокин В. А., Васильев П. П., Пронин С. В. Феномен инсайта и цифровая визуализация активности головного мозга человека // Вестник ЮУрГУ. Серия «Психология». 2017. Т. 10. № 4. С. 47–55.

Шелепин Ю. Е., Фокин В. А., Меньшикова С. В., Борачук О. В., Коскин С. А., Соколов А. В., Пронин С. В., Хараузов А. К.,. Васильев П.П, Вахрамеева О. А. Методы иконики и методы картирования мозга в оценке функционального состояния зрительной системы // Сенсорные системы. 2014. Т. 28. № 2. С. 63–78.

Ardila A., Bertolucci P.H., Braga L.W., Castro-Caldas A., Judd T., Kosmidis M.H., Matute E., Nitrini R., Ostrosky-Solis F., Rosselli M. Illiteracy: the neuropsychology of cognition without reading // Archives of Clinical Neuropsychology. 2010. V. 25. P. 689–712.

Field D.J., Hayes A., Hess R. F. Contour integration by the human visual system: evidence for a local 'association field' // Visual Research. 1993. V. 33. P. 173–93.

Foreman N., Hemmings R. The Gollin incomplete figures test: a flexible, computerised version // Perception. 1987. V. 4. № 16. P. 543.

Hess R.F., Field D.J. Integration of contours: new Insights // Trends in Cognitive Sciences. 1999.V. 3. № 12. P. 480–486.

Kounios J., Beeman M. The cognitive neuroscience of insight // Annual Rev. Psychology. 2014. V. 65. № 93. P. 74.

Kraft A., Grimsen C., Kehrer S., Bahnemann M., Spang K., Prass M., Irlbacher K., Köhnlein M., Lipfert A., Brunner F., Kastrup A., Fahle M., Brandt S.A. Neurological and neuropsychological characteristics of occipital, occipito-temporal and occipito-parietal infarction // Cortex; a Journal Devoted to the Study of the Nervous System and Behavior. 2006. V. 56. P. 38–50.

Saaty T.L. The analytic hierarchy and analytic network measurement processes: Applications to decisions under Risk // European Journal of Pure and Applied Mathematics. 2008. V. 1. № 1. P. 122–196.

Wechsler H. Neural networks for perception: human and machine perception. Academic Press, NY, 2014. 384 p.

Weng J., Paslaski S., Daly J., Van Dam C., Brown J. Modulation for emergent networks: Serotonin and dopamine // Neural Networks. 2013. V. 41. P. 225–239.

Глава 10.

Взаимодействие крупномасштабных нейронных сетей при распознавании текстур

А. К. Хараузов, П. П. Васильев, Ю. Е. Шелепин, А. В. Соколов, В. А. Фокин

Современное представление о функционировании мозга базируется на сетевом подходе, в котором удаленные друг от друга области мозга объединяют на основе их синхронной активности и участии в выполнении одинаковых функций. В настоящее время, на основе данных функциональной магнитно-резонансной томографии (фМРТ), выделяют несколько крупномасштабных нейронных сетей. Например, так называемая «центральная исполнительная сеть» (central executive network, CEN), объединяющая области дорсолатеральной префронтальной коры и задней теменной коры. Активность этой сети связывают с функционированием оперативной памяти и принятием решений при целенаправленном поведении. Области передней цингулярной извилины и переднюю часть островковой зоны мозга объединяют в другую сеть (salience network, SN), отвечающую за обнаружение и переключение внимания на значимые сигналы, поступающие из внешнего мира и происходящие внутри организма. Выделяют также нейронные сети, отвечающие за функции памяти, способности к языку, обработку информации от органов чувств и так далее [см. обзоры Menon, 2011; Nekovarova et al., 2014]. Общим свойством всех нейронных сетей мозга является внутренняя функциональная связь и оппонентные взаимодействия с так называемой сетью пассивного режима (default mode network), которая, по всей видимости, координирует их работу.

Сеть пассивного режима работы мозга была впервые описана в 2001 году Маркусом Рейчлом с соавторами, которые опубликовали доказательства существования нейронной сети, обладающей необычными свойствами — ее активность снижалась во время выполнения человеком различных заданий и увеличивалась в состоянии покоя. Предположив, что наблюдаемая активность характеризует

состояние бездействия, но в готовности к выполнению задания, они назвали ее Default Mode Network, DMN [Raichle et al., 2001], что дословно переводится на русский язык как «сеть работы в режиме по умолчанию» (иногда название переводят как сеть базового режима). В состав сети пассивного режима входят удаленные друг от друга и, на первый взгляд, разобщенные области мозга: медиальная префронтальная кора, задняя поясная извилина, преклиновидная кора и медиальная, латеральная и нижняя части теменной коры. Однако между собой эти области имеют прочную функциональную связь за счет когерентной активности [Fox and Raichle, 2007].

С момента первой публикации данных о сети пассивного режима работы мозга количество работ, посвященных этой теме, увеличивалось лавинообразно. К настоящему времени эта концепция стала центральной темой в когнитивных нейронауках. Оказалось, что снижение активности сети пассивного режима не зависит ни от модальности стимула, ни от типа решаемых задач, но зависит от степени вовлеченности испытуемого в выполнение задачи или от сложности теста. Одновременно со снижением активности сети пассивного режима наблюдается активация областей мозга, ответственных за выполнение экспериментальной задачи. Поэтому в 2006-м году в журнале Science Маркус Рейчл высказал предположение, что обнаруженная им сеть может играть решающую роль в синхронизации работы всех отделов мозга для того, чтобы они, «как спортсмены на соревнованиях, были готовы начать гонку, как только раздастся выстрел стартового пистолета» [Raichle, 2006].

Помимо сети пассивного режима, во время решения человеком различных задач, дополнительно могут снижать свою активность области мозга, отвечающие за обработку сенсорных сигналов других модальностей, менее задействованных в экспериментальных условиях. Так, в наших исследованиях [Хараузов и др., 2015] мы наблюдали деактивацию в первичных слуховых областях при восприятии фрагментированных изображений лиц, замаскированных динамической зрительной помехой. Такое явление объяснимо с точки зрения экономии ресурсов мозга за счет реципрокного торможения на уровне крупномасштабных нейронных сетей. Активация областей мозга, занятых в обработке сигнала, тормозит активность в областях, не задействованных в выполнении конкретной задачи.

Таким образом, соблюдается принцип оппонентности в работе крупномасштабных нейронных сетей головного мозга, выявляемый в основном в томографических исследованиях. Несмотря на то, что метод фМРТ основан на анализе различий в потреблении кислорода тканями, имеются веские доказательства [см. обзор Singh, 2012], что он позволяет оценивать суммарное изменение активности нейронов в изучаемых областях мозга. Поэтому, данный метод является единственным неинвазивным способом, позволяющим анализировать увеличение или снижение активности на уровне крупномасштабных нейронных сетей головного мозга.

Обычно в томографических исследованиях основное внимание уделяется поиску областей мозга, увеличивающих свою активность в процессе выполнения человеком различных задач. Однако для того, чтобы понимать принципы работы мозга как единое целое, необходимо учитывать мозаику распределения деактивированных областей, или снижения активности целых нейронных сетей, происходящих одновременно с активацией. Интуитивно понятно, что в едином пространстве мозга средний объем активированных областей должен изменяться пропорционально объему областей, снизивших свою активность. Это положение и было доказано нами в фМРТ исследованиях при распознавании человеком текстур различной сложности.

Исследования проводили с помощью набора тестов, разработанных в Институте физиологии им. И. П. Павлова РАН, для электрофизиологических и томографических исследований на человеке в задачах оценки статистических свойств сложных изображений [Хараузов и др., 2006; Шелепин и др., 2009; Harauzov et al., 2016]. Тесты представляют набор матриц, каждая из которых состоит из 100 решеток Габора разной ориентации (от 0 до 165 градусов с шагом в 15 градусов). Матрицы различаются по количеству однонаправленных элементов, другими словами, по степени упорядоченности. Доминирующая ориентация в матрицах может быть либо вертикальной, либо горизонтальной. В более упорядоченных матрицах, с большим количеством вертикальных (или горизонтальных) элементов, вероятность появления решеток с другими ориентациями равномерно снижалась. Поскольку всего в каждом изображении было по 100

элементов, при обозначении матриц мы использовали процентное содержание однонаправленных решеток (рис. 1).

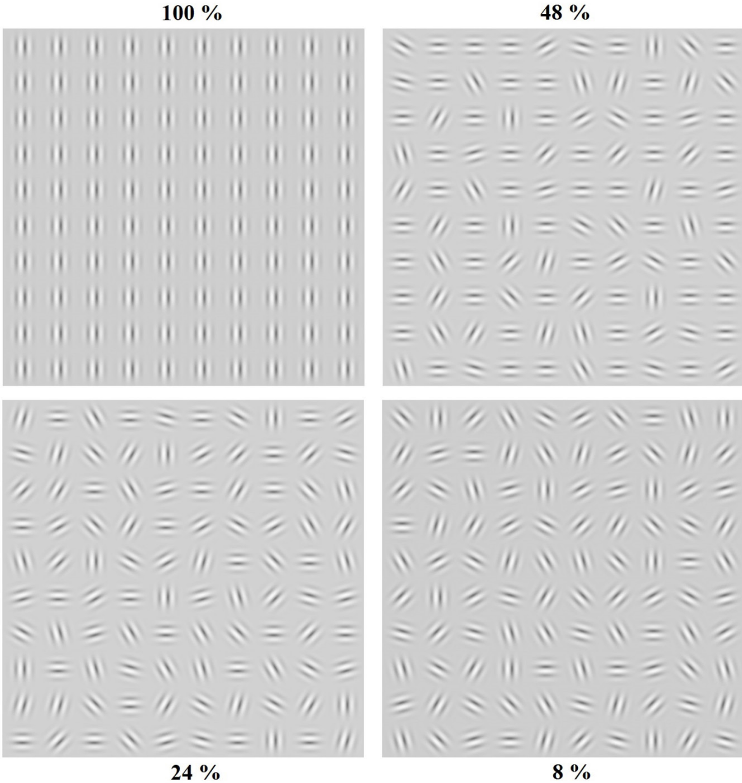


Рис. 1. Примеры тестовых матриц различной степени упорядоченности. Цифры у каждой матрицы обозначают относительное количество вертикальных или горизонтальных элементов. Задача — определить доминирующую ориентацию

Задачей испытуемых было определить, каких элементов в изображении больше — вертикальных или горизонтальных. В данной работе использовали 4 типа матриц, содержащих 100, 48, 24 и 8 однонаправленных элементов. В максимально упорядоченном стимуле все 100 решеток имели одну ориентацию, выделить доминирующую ориентацию в таком стимуле не представляло труда. В минимально упорядоченном стимуле каждая из 12-и ориентаций была представ-

лена в матрице равновероятно, количество решеток каждой ориентации составляло 8. Как видно из рис. 1, по мере уменьшения количества однонаправленных элементов, задача категоризации матрицы и отнесению ее к «вертикальному» или «горизонтальному» типу усложняется. Преимущество данного набора стимулов в том, что изменение сложности можно легко контролировать, при этом основные физические характеристики стимулов, такие как яркость, контраст, размер и т. п. остаются постоянными.

В исследовании принимало участие 25 испытуемых, с остротой зрения равной единице, без неврологических заболеваний. На рис. 2 представлены результаты измерений ответной реакции испытуемых — количества правильных ответов и времени реакции. Наибольшее затруднение вызывали матрицы, содержащие 8 % однонаправленных элементов — медиана времени реакции составила 1080 мс. Поскольку все ориентации в таких матрицах представлены равновероятно, дать правильный ответ в принципе невозможно. По сути, отвечая на матрицы 8 %, испытуемые совершали случайное гадание. По мере увеличения степени упорядоченности стимула среднее количество правильных ответов увеличивается, выходя на плато в ответ на матрицы 100 %. Медиана времени реакции, наоборот, линейно снижается с увеличением количества однонаправленных элементов. Представленные данные ответной реакции наглядно демонстрируют, что по мере увеличения упорядоченности стимула сложность задачи снижается.

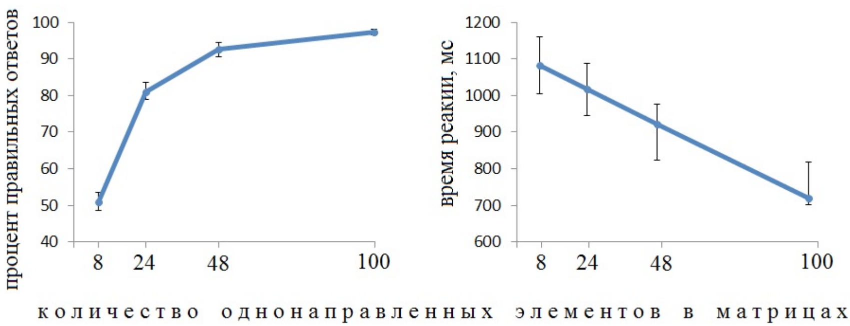


Рис. 2. Зависимость эффективности распознавания и времени реакции от степени упорядоченности стимула

Во время выполнения тестов у испытуемых регистрировали изменение активности мозга с помощью метода функциональной магнитно-резонансной томографии по стандартной методике измерения уровня оксигенации гемоглобина—BOLD (Blood Oxygenation Level Dependent). Различие магнитных свойств оксигемоглобина и дезоксигемоглобина позволяет локализовать наиболее активные участки мозга в момент стимуляции по отношению к состоянию относительного «покоя». В наших исследованиях активность мозга во время распознавания матриц сравнивали с состоянием бездействия, длившегося 5 минут, в котором единственной задачей испытуемым было смотреть на точку в центре экрана. Во время стимуляции, матрицы каждой из четырех степеней упорядоченности предъявляли с частотой 1 Гц блоками по 30 секунд. Последовательность блоков была случайной. В общей сложности каждый тип матриц предъявляли по 5 минут.

На рис. 3 представлены карты распределения активности при решении задач различной сложности на основе усреднения данных по 25-и испытуемым. Красным цветом обозначены области мозга, статистически значимо увеличивающие свою активность относительно состояния покоя. Синим цветом обозначены области мозга, достоверно снижающие активность во время выполнения задания. За критерий достоверности был выбран порог $P<0.01$. Как следует из рис. 3, общий паттерн активированных и деактивированных областей не претерпевает существенных изменений с увеличением сложности задачи. Состав и локализация областей практически не изменяются, но увеличивается интенсивность изменения BOLD-сигнала и объем активированных (и деактивированных) областей мозга при решении более сложных задач.

Среди областей мозга, увеличивающих свою активность во время выполнения тестов, наиболее объемной является затылочно-височная область, включающая зону 19 по Бродману и фузиформную извилину. Хорошо известно, что эти области мозга ответственны за восприятие сложных изображений. Вторая по объему активация затрагивает области в верхней части теменной коры, ответственные за локализацию объектов в пространстве. В лобных долях мозга активировались участки средней лобной извилины и дополнительной

моторной области. В подкорковых структурах была выявлена билатеральная активация в ядрах подушки таламуса.

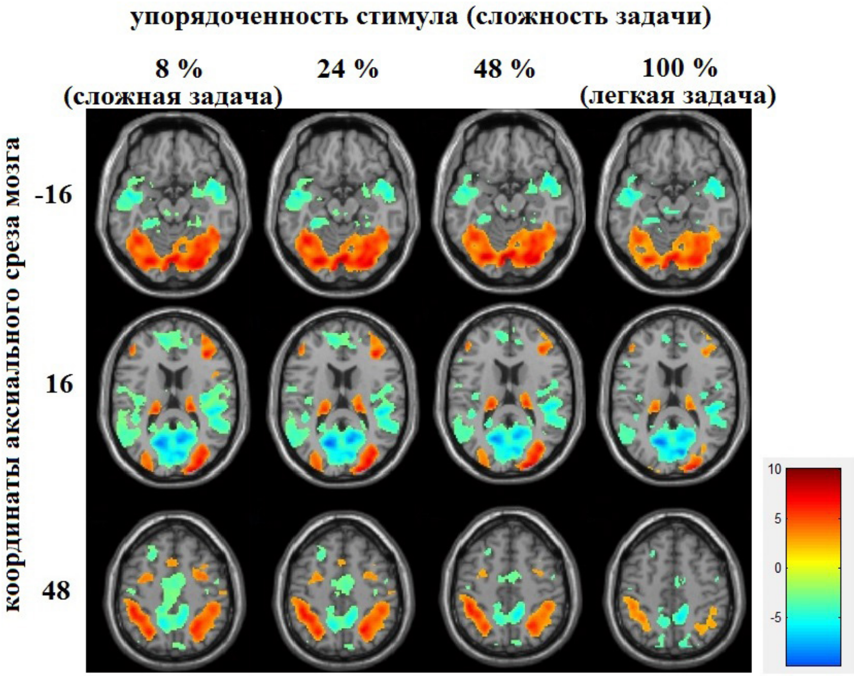


Рис. 3. Распределение активности мозга во время выполнения испытуемыми задач распознавания различной сложности. Красным цветом обозначены области, увеличивающие активность. Синим цветом обозначены области, снизившие активность относительно состояния покоя. Усреднены данные 25-и испытуемых, P<0.01

В целом, карта активации в наших исследованиях совпадает с результатами, полученными во многих лабораториях мира в экспериментах, связанных с принятием решений при восприятии изображений [Heekeren et al., 2004; обзор — Keuken et al., 2014]. Причем, как показывают наши ранние исследования, активация в лобных областях отражает работу механизмов принятия решения о типе стимула и реализации ответной реакции [Шелепин и др., 2009]. Это утверждение было доказано путем сопоставления карт активации,

полученных при пассивном наблюдении предъявляемых стимулов и во время выполнения задачи по распознаванию доминирующей ориентации в матрицах. При пассивном наблюдении у испытуемых наблюдали активацию только в затылочно-височных и теменных областях мозга. Когда испытуемые выполняли задачу по распознаванию, у них дополнительно активировались участки мозга в лобных долях.

Наряду с увеличением активности, в ряде областей наблюдали значимое снижение уровня BOLD-сигнала относительно состояния покоя. Среди них были обнаружены структуры, входящие в состав сети пассивного режима работы мозга: задняя поясная извилина, медиальная префронтальная кора, преклиновидная кора и части теменной коры. Как уже упоминалось, для этих областей характерно снижение активности во время целенаправленной деятельности, вне зависимости от модальности стимула. Дополнительно, наблюдали значимую деактивацию в верхней височной извилине, в которой, как известно, находятся области, ответственные за переработку слуховой информации. Таким образом, подтверждаются наши ранние данные о том, что при определенных типах деятельности, помимо компонентов сети пассивного режима, могут снижать активность области мозга, менее задействованные в выполнении задания [Хараузов и др., 2015].

Для того чтобы оценить, насколько сбалансированы процессы активации и деактивации в головном мозге, мы рассчитали и сопоставили общий объем всех активированных и деактивированных областей при решении задач разной сложности. У каждого испытуемого, в ответ на каждый тип матриц, отдельно подсчитывали количество вокселей, со значимым увеличением и уменьшением BOLD-сигнала относительно состояния покоя. Размер воксела (минимального элемента объемного изображения мозга) составлял 2×2×2 мм. Полученные результаты затем усреднили по всем 25-и испытуемым и построили зависимость количества активированных и деактивированных вокселей, от сложности задачи.

Результаты расчетов представлены на рис. 4. Суммарный объем областей мозга, увеличивших свою активность в процессе выполнения тестов, оказался равен объему областей, в которых было зафиксировано снижение активности. Совпадение объемов наблю-

дали при решении задач всех четырех уровней сложности. По мере упрощения задачи количество активированных и деактивированных вокселей пропорционально снижалось, но их соотношение всегда оставалось постоянным. Можно заключить, что процессы активации и деактивации по всему объему мозга сбалансированы — чем сильнее активируются одни области, тем сильнее снижается активность в других областях.

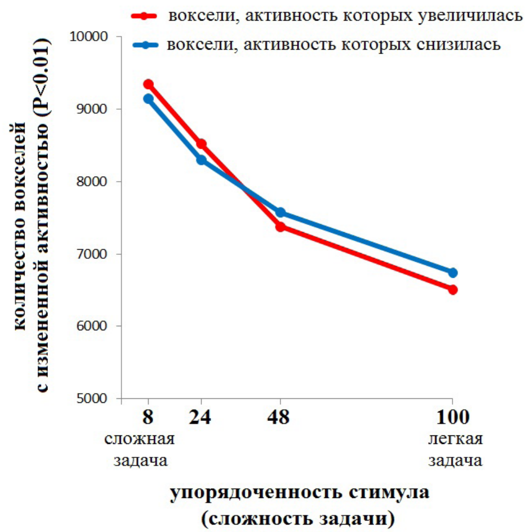


Рис. 4. Зависимость суммарного количества вокселей с измененной активностью относительно состояния покоя, от сложности задачи. Красным цветом обозначены воксели, увеличивающие активность. Синим цветом обозначены воксели, снизившие активность. Усреднены данные 25-и испытуемых, $P<0.01$

Таким образом, при выполнении задач, связанных с распознаванием и принятием решений, в мозге человека последовательно активируется несколько различных нейронных механизмов. В первую очередь вступают в работу нейронные сети, обеспечивающие обработку сенсорной информации. Затем должны быть задействованы гипотетические нейронные механизмы, обеспечивающие сравнение поступающего сигнала с внутренним шаблоном (или эталоном), хранящимся в памяти. После чего происходят процессы принятия решений, планирования и реализации моторной реакции. Эти и другие нейронные

сети, которые могут принимать участие в обеспечении правильного ответа, увеличивают свою активность относительно состояния покоя.

Параллельно с активацией, согласно современным представлениям, снижается активность в нейронной сети пассивного режима работы мозга. Кроме того, согласно нашим данным, снижается активность нейронных сетей, менее задействованных в выполнении задания. Так, при распознавании зрительных стимулов, снижается активность в слуховых областях мозга. То есть, происходит реципрокное торможение на уровне крупномасштабных нейронных сетей. Причем, как показывают наши данные, имеется тенденция к поддержанию постоянства среднего уровня нейронной активности. Вероятно, это обеспечивается путем автоматического перераспределения активности между различными областями мозга. Более подробно этот вопрос рассмотрен в статье [Хараузов и др., 2018].

Для понимания механизмов поддержания баланса активности необходимо исследовать временную динамику взаимодействия крупномасштабных нейронных сетей. Временное разрешение метода фМРТ измеряется секундами. Понятно, что при такой полосе пропускания теряется важная информация о высокочастотной динамике нейронной активности. Перспективным путем исследования является комбинация томографических и электрофизиологических измерений, так как временное разрешение электроэнцефалограммы измеряется миллисекундами. Так, на примере распознавания доминирующей ориентации в матрицах, в электрофизиологических исследованиях на человеке, нами была установлена временная последовательность развития некоторых нейронных процессов, задействованных в выполнении данной задачи [Harauzov et al., 2016]. В частности, было показано, что механизмы принятия решения о типе стимула начинают активироваться через 400 мс после предъявления стимула.

К сожалению, неинвазивные электрофизиологические методы также имеют свои недостатки — низкое пространственное разрешение и искажение сигнала костными, мышечными и другими тканями. В литературе имеются немногочисленные данные по интракраниальной регистрации у человека нейронной активности в непосредственной близости от структур мозга, функционирующих в противофазе. Однако, несмотря на высокую информативность данного

подхода, метод интракраниальной регистрации нейронной активности у человека сопряжен с рядом этических и технических проблем. В частности, локализация электродов продиктована медицинскими целями и нет возможности охватить все интересующие области. К тому же, исследование проводится на больных, и отделить патологическую активность от нормальной довольно сложно.

Пролить свет на проблему динамики взаимодействия крупномасштабных нейронных сетей помогут, на наш взгляд, инвазивные электрофизиологические исследования на обезьянах. Так, в экспериментах на макаках было показано снижение нейронной активности в задней цингулярной извилине (основном компоненте сети пассивного режима у человека) во время ритмической зрительной стимуляции. При этом активность в зрительных областях мозга предсказуемо повышалась [Bentley, 2016]. Изменение активности коррелировало с уровнем оксигенации крови, регистрируемым инвазивно в тех же структурах. Эти результаты являются, во-первых, прямым подтверждением зависимости BOLD-сигнала, регистрируемого в фМРТ, от нейронной активности. Во-вторых, доказывают существование у обезьян нейронной сети, аналогичной сети пассивного режима у человека, находящейся в оппонентных отношениях с другими сетями мозга. Примечательно, что перед снижением нейронной активности в задней цингулярной извилине в ответ на зрительную стимуляцию, наблюдали кратковременное, но мощное увеличение активности в этой области. Подобные резкие изменения активности в начале стимуляции наблюдали и у человека [Fox et al., 2005; Dosenbach et al., 2006]. Таким образом, открытие у обезьян нейронной сети, аналогичной сети пассивного режима у человека [см. обзор Mantini et al., 2001], открывает широкие перспективы для изучения принципов оппонентных отношений на уровне крупномасштабных нейронных сетей путем непосредственной регистрации нейронной активности.

Литература

Хараузов А. К., Васильев П. П., Соколов А. В., Фокин В. А., Шелепин Ю. Е. // Анализ изображений функциональной магнитно-резонансной томографии головного мозга человека в задачах

распознавания текстур // Оптич. журн. — 2018. — Т. 85, № 8, спец. вып. «Нейроиконика». — С. 22–28 // DOI:10.17586/1023–5086–2018–85–08–22–28

Хараузов А. К., Васильев П. П., Соколов А. В., Шелепин Ю. Е., Кувалдина М. Б., Борачук О. В., Фокин В. А., Пронин С. В. // Восприятие изображений в задачах зрительного поиска в условиях динамической помехи. // Оптический журнал. 2015. Т. 82. № 5. С. 42–55.

Хараузов А. К., Шелепин Ю. Е., Пронин С. В., Сельченкова Т. В., Носков Я. А. // Электрофизиологические исследования механизмов распознавания текстур. // Российский физиологический журнал им. И. М. Сеченова, 2007, Т. 93. № 1. с. 3–13.

Шелепин Ю. Е., Фокин В. А., Хараузов А. К., Пронин С. В., Чихман В. Н. // Локализация центра принятия решений при восприятии формы зрительных стимулов // Доклады Академии Наук, 2009, том 429, № 6, с. 1–3.

Dosenbach N. U. F., Visscher K. M., Palmer E. D., Miezin F. M., Wenger K. K., Kang H. C., Burgund E. D., Grimes A. L., Schlaggar B. L., Petersen S. E. // A core system for the implementation of task sets. // Neuron. 2006. 50:799–812.

Fox, M.D. and Raichle, M.E. // Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. // Nat. Rev. Neurosci. 2007. 8, 700–711.

Fox M. D., Snyder A. Z., Barch D. M., Gusnard D. A., Raichle M. E. // Transient BOLD responses at block transitions. // NeuroImage. 2005. 28: 956–966.

Harauzov A. K., Shelepin Y. E., Noskov Y. A., Vasilev P. P., Foreman N. P. // The time course of pattern discrimination in the human brain // Vision Research 125 (2016) 55–63.

Heekeren H. R., Marrett S., Bandettini P. A., Ungerleider L. G. // A general mechanism for perceptual decision-making in the human brain // Nature, Vol. 431, 14 October, 2004, pp.: 859–861.

Keuken M. C., Müller-Axt C., Langner R., Eickhoff C. B., Forstmann B. U., Neumann J. // Brain networks of perceptual decision-making: an fMRI ALE meta-analysis // Frontiers in Human Neuroscience, June 2014, Volume 8, Article 445, pp.: 1–14.

Mantini D., Gerits A., Nelissen K., Durand J-B., Joly O., Simone L., Sawamura H., Wardak C., Orban G. A., Buckner R. L., Vanduffel W. // Default Mode of Brain Function in Monkeys // The Journal of Neuroscience, September 7, 2011. 31(36):12954–12962.

Menon V. // Large-scale brain networks and psychopathology: a unifying triple network model. // Trends Cogn Sci. Oct. 2011; 15(10):483–506.

Nekovarova T., Fajnerova I., Horacek J. and **Spaniel F.** // Bridging disparate symptoms of schizophrenia: a triple network dysfunction theory. Front Behav Neurosci.; 2014. 8: 171.

Raichle M. E., MacLeod A. M., Snyder A. Z., Powers W. J., Gusnard D. A., Shulman G. L. // A default mode of brain function. // Proc. Natl. Acad. Sci. U.S.A. 2001. 98: 676–682.

Raichle, M. // Neuroscience. The brain's dark energy. // Science 2006. 314, 1249–1250.

Singh K. D. // Which "neural activity" do you mean? fMRI, MEG, oscillations and neurotransmitters. // NeuroImage. 2012. 62:1121–1130.

Bentley W. J., Li J. M., Snyder A. Z., Raichle M. E., Snyder L. H. // Oxygen Level and LFP in Task-Positive and Task-Negative Areas: Bridging BOLD fMRI and Electrophysiology // Cerebral Cortex, January 2016. 26: 346–357.

Глава 11.
О восприятии движения реальной и воображаемой сцены

А. М. Ламминпия, Е. Ю. Малахова

Любое движение представляет собой совершающийся во времени процесс перемещения движущегося тела в пространстве. В большинстве видов деятельности человека большое значение имеет правильное восприятие движений как совершаемых другими людьми, так и собственных.

Сам по себе окуломоторный акт выражает не просто ответную реакцию на проксимальный стимул или выработанный навык. Он является выражением активности субъекта восприятия, направляемой как прошлым и настоящим опытом, так и будущим: определенными намерениями, целями или планами. Мозг человека, его нейронные сети запускают произвольные и непроизвольные движения глаз, следящие за движущимся предметом в окружающем пространстве.

Управление движением глаз зависит от физических и семантических характеристик наблюдаемой сцены, мотивов у наблюдателя или получаемых им инструкций, например, от экспериментатора.

Запуск глаз, прослеживающих движение, сопровождается саккадическим подавлением восприятия, а в момент фиксации взора происходит считывание информации. При чтении прослеживающие движения глаз являются выученными и произвольными. Захват элементов текста происходит в момент фиксации взора и микродвижения глаз.

В работе зрительного восприятия большое значение отводится интеллектуальным механизмам обработки информации. В зависимости от задач, которые стоят перед человеком, то есть в зависимости от характера сведений, которые он должен получить, будет соответственно изменяться и распределение точек фиксации на объекте, поскольку различные сведения обычно локализованы в различных частях объекта.

Имманентный подход подразумевает отношение к тексту как к автономной реальности, нацеленность на выявление его внутрен-

ней структуры, поэтому художественный текст может рассматриваться как вид внутренней «виртуальной» реальности читателя. Таким образом, сюжет текста может быть рассмотрен как движение воображаемой читателем сцены.

Аналогом движения в тексте могут служить глаголы, отражающие динамическую составляющую информации. Поэтому изучение общих характеристик движений глаз при восприятии реального и воображаемого движения представляет интерес.

Эффективным инструментом для изучения перцептивных и когнитивных функций человека является регистрация движения глаз. В.А. Барабанщиков и В.И. Белопольский показали новые пути анализа механизмов, сохраняющих стабильность восприятия окружающего мира при движениях глаз [Барабанщиков, Белопольский, 2008]. В качестве единицы анализа ими был выбран фиксационный поворот глаз, который интерпретируется как функциональная структура, которая подчиняется действию как внешних, так и внутренних детерминант, таких как: прогнозирование конечного результата, способ управления движениями глаз, ведущий уровень на котором они строятся, сопряженность окуломоторики с другими двигательными актами наблюдателя. В центре внимания оказываются устойчивые конфигурации — паттерны — движений глаз, реализующие познавательное (перцептивное) отношение индивида к среде. Они не могут быть сведены к сумме отдельных дрейфов и саккад: каждое из выполняемых движений может быть понято в рамках всего паттерна в целом.

В исследовании восприятия динамической сцены были проанализированы движения глаз во время просмотра классической и модифицированной задачи, индуцирующей слепоту по невниманию, с различными инструкциями [Shelepin et al., 2012]. Первая инструкция — подсчитать количество передач мяча одной из команд [Chabris & Simons, 2010], вторая — заключалась в обычном просмотре видео [Shelepin et al., 2012]. При выполнении первого задания наблюдатели, занятые подсчетом, не замечали передвигающуюся через все поле зрения танцующую фигуру «гориллы». При втором задании не было специальной инструкции для концентрации внимания на деталях. Поэтому испытуемые сразу замечали появление необычной фигуры «гориллы» в игровом поле. На рис. 1 (два верхних кадра)

кружками показаны единичные фиксации при выполнении первого задания. Взор фиксирован на руках игроков в белых майках и на мяче. При выполнении второго задания взор фиксирован на необычной фигуре (в кадрах, изображенных внизу). Инструкции определяли разный тип саккадических движений глаз при просмотре одного и того же сюжета. На рис. 2 представлены тепловые карты (суммарные значения фиксаций в разных точках поля зрения) за время просмотра ролика. Видно, что первое задание действительно организует движения глаз в полном соответствии с инструкцией — слежением за мячом и руками игроков в белых майках. Второе задание организует другой тип движений глаз по экрану, инструкции просто смотреть видео. Этот тип движений глаз характерен редкими саккадами и длительными фиксациями в центре экрана [Шелепин, 2017]. «Планирование и запуск» саккад определяет крупноклеточная, магноцеллюлярная нейронная сеть, а «считывание» информации во время фиксации и тремора глаз — мелкоклеточная парвоцеллюлярная нейронная сеть. Инструкции определяли саккадические движения глаз



Рис. 1. Примеры одиночных кадров взяты из работы[Chabris & Simons, 2010]. В каждом кадре маленький кружочек — пример одиночной фиксация взора при разных инструкциях [Shelepin et al., 2012].

с различными параметрами: в первом случае паттерн движения глаз был распределен по всей площади сцены, что характерно для активного просмотра и решения задачи, а во втором — большее количество фиксации было сосредоточено в центре сцены, как при пассивном просмотре (рис. 2). Также варьировали угловой размер тестовых изображений для анализа значимости области восприятия во время фиксаций глаз. Площадь невнимательности зависит от инструкции и углового размера тестовых изображений.

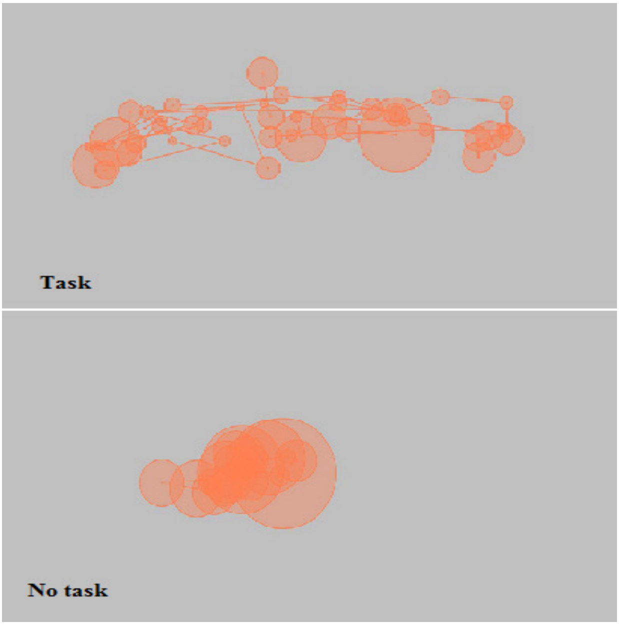


Рис. 2. Паттерны движений глаз при выполнении задачи подсчета (*Task)* и при простом просмотре *(No task)*.

На основе полученных данных был сделан важный вывод [Shelepin et al., 2012] о том, что эффект слепоты по невниманию в его классической динамической парадигме можно объяснить не только механизмами внимания, но, главное, хорошо известным эффектом саккадического подавления [Burr, Ross, Morrone, 1994]. Поскольку саккадическое подавление связано с магноцеллюлярной нейронной сетью, отвечающей и за детекцию движения, и за высо-

кочастотное, временное описание наблюдаемой сцены, известно, что роль этой системы проявится и при осознании воображаемого движения, а именно, при чтении [Шелепин, 2017].

Так как в тексте аналогом движения могут служить глаголы, изучение роли частей речи в языке также затрагивает оценку их важности для передачи, понимания и восприятия информации. При помощи эксперимента с перестановкой букв внутри слова [Furtner, Rauthmann, Sachse, 2009] было показана роль существительных, как содержащих большую часть информации и поэтому сделан вывод о том, что они могут быть рассмотрены как наиболее значимые для понимания и являющиеся основной семантической составляющей текста. Исследователями была высказана гипотеза о том, что слова, на которые приходится большее количество фиксаций, играют более важную роль для понимания. По результатам их эксперимента в 49 % всех случаев повторных фиксаций для повышения понимания сложных транспонированных слов читатели обращались к существительным. При этом в тексте содержалось около 26 % существительных. Прирост в количестве повторных фиксаций, приходящихся на прилагательные, составил порядка 8 %, в глаголах значимой разницы обнаружено не было. Слова из закрытых классов (местоимения, предлоги, артикли, союзы, междометия) использовались испытуемыми для повышения понимания текста в 16 % всех случаев, при этом их относительная частота в тексте составляла 42 %. Эти исследования показывают, что на параметры чтения оказывают влияние синтаксические категории, такие как часть речи и принадлежность к закрытой либо открытой синтаксической группе. В начале 80-х годов Карпентер и Джаст показали, что в процессе чтения фиксация не осуществляется на каждом слове, кроме того синтаксические и семантические характеристики отдельного слова оказывают влияние на то происходит фиксация или не происходит: читатели в ходе исследования в среднем фиксировали 67,8 % слов читаемого текста, причём на фиксацию значимых частей речи приходилось 83 % отведенного времени, а на фиксацию служебных частей речи приходилось только 38 % времени. [Just, Carpenter, 1980]. Кроме того, количество и длительность фиксаций коррелирует с длиной слова и его частотой в языке [Rayner, 1998]. Так, фикса-

ции на синтаксических категориях, относящихся к открытому классу (существительные, глаголы, прилагательные, наречия), являются более продолжительными по сравнению с фиксациями на синтаксических категориях, относящихся к закрытому классу [Trauzettel-Klosinski, Dietz, 2012].

Результаты последних исследований показывают, что параметры фиксации взгляда позволяют различать большинство синтаксических категорий в английском языке и аннотировать корпус в случае заранее заданного словаря тегов для разметки потенциальных категорий [Barrett, Sogaard, 2015]. В качестве тегов для разметки применялся набор из двенадцати универсальных категорий частей речи [Petrov, Das, McDonald, 2011].

Перемещение центральной области, обладающей наивысшей остротой зрения, к интересующему объекту осуществляется посредством саккад. Во время выполнения саккады зрительная информация не воспринимается (эффект саккадического подавления) [Erdmann, Dodge, 1920], некоторые когнитивные функции также подавляются [Irwin, Carlson-Radvansky, 1996]. Однако, ряд работ [Irwin, 1998] указывает на осуществление лексической обработки в процессе совершения саккады, что подтверждает необходимость изучения характеристик саккад при исследовании восприятия текстовой информации. Большинство работ, посвященных чтению, фокусируются на исследовании частоты и продолжительности фиксаций как непосредственно отражающих процесс осмысления текста [Rayner,1998].

В исследовании влияния изменения частеречного состава текста (а именно увеличения числа глаголов) мы рассмотрели влияние синтаксической информации на параметры движений глаз в процессе чтения русскоязычных текстов носителями языка. В качестве стимулов были использованы текстовые фрагменты классической литературы, в которых содержание одной из частей речи (глаголы, существительные, прилагательные) превышало среднее по тексту значение более чем на три среднеквадратичных отклонения. Регистрация окуломоторной активности проводилась методом eye-tracking с частотой дискретизации в 1000 Гц. Тексты предъявлялись с расстояния 60 см от глаз наблюдателя при помощи LCD-монитора с разрешением экрана 1920x1080. Задачей каждого наблюдателя было прочесть 15

текстовых фрагментов, представляющих три группы в зависимости от частеречного состава. В качестве стимулов использовались фрагменты из произведений классической литературы таких авторов как Толстой Л.Н., Булгаков М.А, Шолохов М.А., Достоевский Ф.М. Для отбора стимульного материала было проанализировано более 800 тысяч фрагментов, содержащих 100 ± 2 слова и состоящих из законченных предложений. Так как для расчета и сопоставления содержания частей речи в текстовых фрагментах учитывалось количество отдельных слов, а не количество символов, длина текстовых фрагментов, выраженная в символах, при этом варьировалась как между отдельно взятыми стимулами, так и в среднем по группам. Средняя длина слов также отличалась: для текстов с прилагательными она составляла 7.0 символов, для текстов с существительными — 6.3, для текстов с глаголами — 5.6. Чтобы избежать влияния длины фрагментов на характеристики прочтения, учитывался относительный состав текста, а переменные, напрямую зависящие от длины текста, были нормализованы. Регистрировалось время начала и окончания совершения саккады, ее амплитуда в угловых минутах, длительность, средняя и пиковая скорость, наклон кривой, рассчитываемый как отношение разницы в скоростях к разнице во времени для начальной точки и первого локального максимума, а также длительность следующей за саккадой фиксации. Также были рассчитаны вторичные показатели: количество совершенных для прочтения текста саккад, количество возвратных саккад, отношение количества прямых саккад к возвратным, среднее количество символов, захватываемых за одну прямую саккаду, являющееся более универсальным показателем для измерения длины саккады при чтении текста, чем амплитуда, так как остается неизменным при изменении размера шрифта и расстояния наблюдателя от предъявляемого текста [Morrison, Rayner, 1981]. Использование характеристики количества захватываемых за одну саккаду символов также позволяло избежать влияния различной длины текстов на последующий анализ. Кроме того параметры, зависящие от длины текста (общее время прочтения и фиксаций, количество прямых и возвратных саккад), были приведены к относительным показателям путем нормирования на количество символов в текстовом фрагменте.

Так как перед испытуемыми стояла задача естественного чтения,

т.е. оптимального индивидуального способа работы с текстом, это усиливало индивидуальные различия в группе испытуемых и усложняло поиск общих закономерностей процесса чтения. На рис. 3 приведено время, затрачиваемое наблюдателем при прочтении одного текстового фрагмента (Рис. 3). Высокая вариативность показателей среди участников согласуется с ранее полученными данными о дисперсии в скорости чтения для носителей русского языка, составляющей 92 % между испытуемыми и 8 % для одного наблюдателя между различными текстами [Trauzettel-Klosinski, Dietz, 2012].

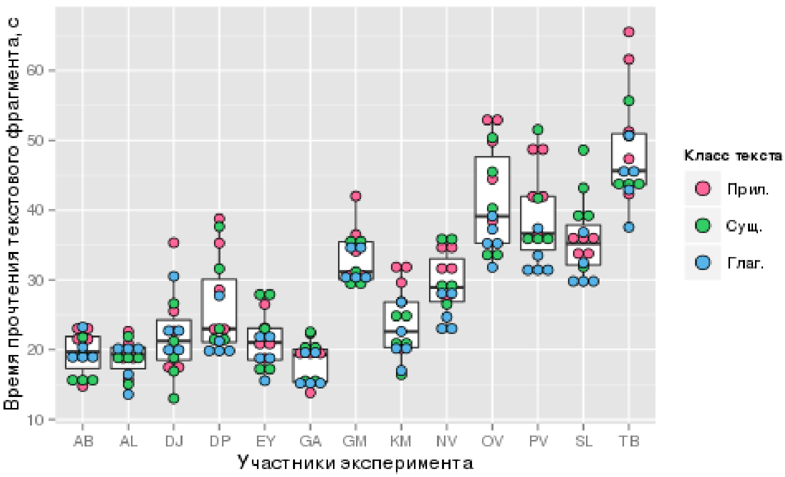


Рис. 3. Время прочтения одного текстового фрагмента у испытуемых

Среднее время прочтения для наблюдателей варьировалось в диапазоне от 19 до 48 с. Подобный разброс делает нецелесообразным анализ показателей, усредненных по всем участникам. Вследствие чего для каждого испытуемого на основе полного объема мониторинговых данных нами были рассчитаны средние для него значения показателей окуломоторной активности. Также было рассчитаны средние значения по отдельно взятому стимулу вне зависимости от читающего. При обработке информации оценивалось отклонение параметра для конкретного стимула от среднего для испытуемого значения.

На рис. 4 приведены диаграммы размаха, позволившие оценить

характер зависимости между отклонением от среднего параметра окуломоторной активности и группой текста, на которой фиксировались наблюдения. Так, при чтении текстов различных групп существенные изменения в характеристиках чтения наблюдателя по сравнению с его средними показателями были выявлены: в количестве захватываемых символов (среднее значение и стандартное отклонение в категории прил. — 0.25 ± 0.97, сущ. — 0.29 ± 1.79, глаг. — –0.54 ± 1.27 от среднего для испытуемого количества символов); в длительности саккады (прил. — 0.6 ± 1.08, сущ. — –0.5 ± 1.77, глаг. — 0.06 ± 1.42, мс); в средней скорости (прил. — 0.17 ± 0.38, сущ. — –0.2 ± 0.65, глаг. — 0.14 ± 0.67, угл.мин/мс) и наклоне скоростного профиля саккады (прил. — 0.01 ± 0.03, сущ. — –0.02 ± 0.05, глаг. — 0.02 ± 0.06) (Рис. 4, 5).

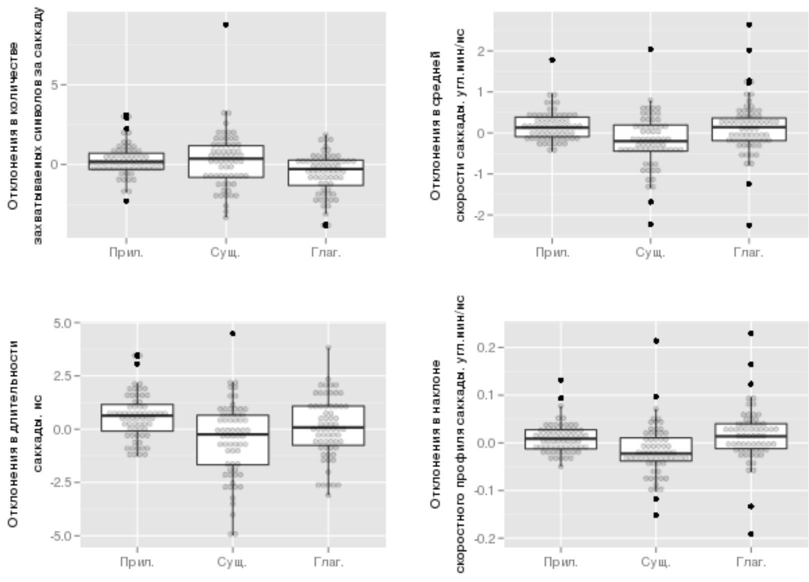


Рис. 4. Точечные диаграммы и диаграммы размаха для характеристик чтения, существенно отличающихся между классами текстов. По оси ординат указаны отклонения данного параметра от среднего для испытуемого значения. По оси абсцисс — группы текстовых фрагментов

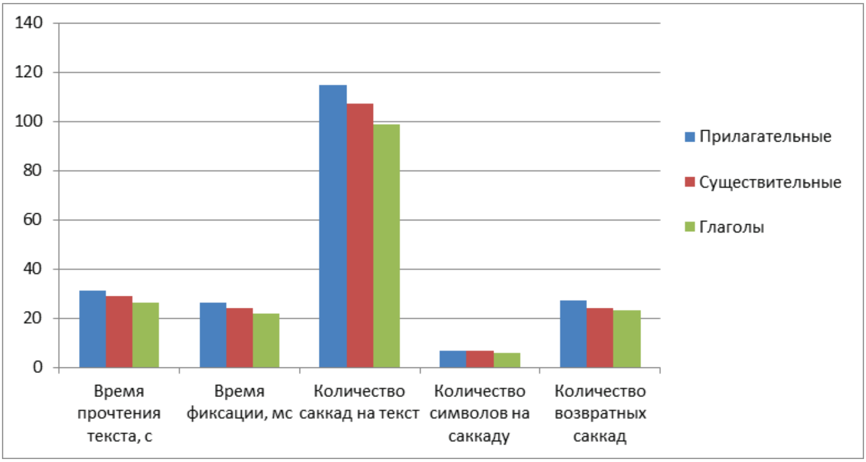


Рис. 5. Основные параметры движений глаз при чтении текстов с преобладанием одной из частей речи

Анализ полученных данных показал, что при чтении текстов, насыщенных глаголами, уменьшается среднее количество захватываемых за одну саккаду символов (Рис. 5). Одной из причин может являться наименьшая среди групп средней длина слова — 5,6 символа. Так как глаголы в текстовых фрагментах представлены наиболее короткими словами, но при этом, не могут быть опущены как второстепенные части речи [Fitzsimmons, Drieghe, 2011], возникает необходимость осуществления большего количества саккад на фиксированное количество символов. Кроме того, окончания глаголов могут нести в себе важную информацию (род, число, время). Было показано, что при наличии существенной информации в окончании слова вероятность совершения второй фиксации на том же слове возрастает [Hyönä, Niemi, 1989]. Процесс чтения текстов с повышенным содержанием существительных характеризуется более низкой средней скоростью саккады и пологим наклоном её кривой по сравнению с двумя другими группами слов, что может быть вызвано снижением динамики повествования из-за уменьшения количества глаголов в целом и глаголов действия, в частности. В целом, паттерн окуломоторной активности при чтении текстов с преобладанием составляющих «движения» — глаголов — можно охарактеризовать как

менее детализированный и более динамичный, по сравнению с чтением других типов текстов.

Таким образом, относительно незначительные изменения реальной или воображаемой сцены (такие как доминирование определенных частей речи в тексте или инструкция для работы с материалом) заметно отражаются на параметрах окуломоторной активности наблюдателя, что может свидетельствовать об изменениях в организации нейронной сети, координирующей и определяющей стратегию наблюдателя.

Литература

Барабанщиков В. А., Белопольский В. И. Стабильность видимого мира. Москва: ИП РАН, 2008. 298 с.

Шелепин Ю. Е. Введение в нейроиконику // «ИТК Троицкий мост», Спб. 2017. ISBN978–5–6040327–1–8

Barrett M., Søgaard A. Reading behavior predicts syntactic categories // CoNLL. 2015. P. 345.

Burr D. C., Morrone M. C., Ross J. Selective suppression of the magnocellular visual pathway during saccadic eye movements // Nature. 1994. N. 371. P. 511–513.

Chabris C., Simons D. The Invisible Gorilla. USA, Crown Publishing Group, 2010. ISBN: 0307459659.

Erdmann B., Dodge R. Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage // In Griffith C. R. (Ed.). Halle: Nie.Esmeyer, 1920.

Furtner M. R., Rauthmann J.F, Sachse P. Nomen est omen: Investigating the dominance of nouns in word comprehension with eye movement analyses // Advances in Cognitive Psychology. 2009. V. 5. P. 91–104.

Hyönä J., Niemi P., Underwood G. Reading long words embedded in sentences: Informativeness of word halves affects eye movements // Journal of Experimental Psychology: Human Perception and Performance. 1989. 15.1. P. 142.

Irwin D.E, Carlson-Radvansky L.A. Cognitive suppression during saccadic eye movements // Psychological Science. 1996. P. 83–88.

Irwin D.E. Lexical processing during saccadic eye movements // Cognitive Psychology. 1998. 36.1. P. 1–27.

Just M.A., Carpenter R.A. A theory of reading: From eye fixations to comprehension // Psychological Review. 1980. V.87. P. 329–354.

Morrison R.E., Rayner K. Saccade size in reading depends upon character spaces and not visual angle // Attention, Perception, & Psychophysics. 1981. 30.4. P. 395–396.

Petrov S., Das D., McDonald R. A universal part-of-speech tagset [Electronic resource] // URL: https://arxiv.org/pdf/1104.2086.pdf. 2011.

Rayner K., Duffy S. On-line comprehension processes and eye movements in reading // Reading research: Advances in theory and practice. New York: Academic Press, 1988. P. 13–66.

Rayner K. Eye movements in reading and information processing: 20 years of research // Psychological Bulletin. 1998. V. 124. P. 372–422.

Shelepin Y., Kuvaldina M., Harauzov A., Vakhrameeva O., Pronin S., Yamschinina P. Investigation of the inattentional blindness for dynamic events as a result of saccadic suppression. // Perception. 2012. V. 41. P. 229.

Trauzettel-Klosinski S., Dietz K., Group I. R. S. Standardized assessment of reading performance: the new International Reading Speed Texts IReST // Investigative ophthalmology & visual science. 2012. 53.9. P. 5452–5461.

Нейротехнологии и искусственные нейронные сети

Глава 12.
Современные нейронные сети глубокого обучения для автоматического анализа изображений

Р. О. Малашин, С. В. Пономарев

Глубокая нейронная сеть — это искусственная нейронная сеть с несколькими скрытыми слоями [Schmidhuber, 2015]. Дополнительные слои позволяют строить абстракции более высоких уровней, что дает возможность формировать внутренние представления (модели) сложных объектов реального мира, которые могут быть использованы для принятия решений в различных задачах. Обычно используются сети прямого распространения, не имеющие обратных связей, однако, показано успешное применение глубоких архитектур и в рекуррентных сетях [Mikolov et al., 2010].

К настоящему времени методы, использующие нейронные сети глубокого обучения, заняли ведущее положение в автоматическом анализе изображений. Наиболее успешные решения этой области ранее базировались преимущественно на аналитических правилах и эвристиках, некоторые из которых оказались чрезвычайно эффективными. Например, неподготовленный человек во многих случаях справляется с задачей сопоставления аэрокосмических снимков хуже, чем система, основанная на автоматическом поиске и сопоставлении ключевых точек, которые выбираются, описываются и сопрягаются по заранее заложенному и достаточно универсальному алгоритму [Malashin, 2014]. Однако возможность создания более сложных систем (а тем более сильного искусственного интеллекта) становится на грань практически (и даже теоретически) неразрешимой проблемы, если эту задачу решать лишь на основе фиксированных правил. Проблема преодолевается путем предоставления возможности алгоритму анализировать информацию «самостоятельно» — обучаться.

Необученная нейронная сеть представляет собой не алгоритм, а лишь спецвычислитель, который состоит из связанных элементарных единиц обработки информации — идеальных нейронов. Такой

вычислитель способен реализовать один из алгоритмов, для которого предназначена его архитектура. Например, структура сверточных нейронных сетей (СНС), разработанная для анализа изображений, была выбрана с учетом знаний о функционировании ганглиозных клеток сетчатки, простых и сложных клеток зрительной коры.

С точки зрения пространства решаемых задач искусственные нейронные сети не имеют преимуществ над другими архитектурами вычислений, однако для них существует эффективный алгоритм настройки свободных параметров (весов связей между нейронами), который дает возможность формировать внутри обученной нейронной сети глубокого обучения сложные экстракторы признаков. Именно в этом кроется причина успеха современных нейронных сетей при решении сложных задач автоматического анализа изображений.

Модель идеального нейрона и обучение глубоких нейронных сетей

Классическую модель идеального нейрона, наиболее эффективно используемую в современных нейронных сетях, можно описать схемой, приведенной на рис. 1,

где: $x_i \in \mathbf{X}$ — отклик нейрона предыдущего слоя;

$w_i \in \mathbf{W}$ — вес связи одного из нейронов предыдущего слоя с одним из нейронов текущего слоя;

b — вес связи с постоянным значением (связь с единицей);

$y = f(X, W, b) \in \mathbf{Y}$ — выход нейрона текущего слоя;

$h = f(y)$ — итоговый отклик нейрона текущего слоя, получается путем применения активационной функции (АФ) f к выходу нейрона.

Значения x_i, w_i, b представляются в виде вещественных чисел. В современных нейронных сетях глубокого обучения для автоматического анализа изображения на первом слое значения x_i кодируют яркости одного из цветовых каналов.

Принято рассчитывать отклик нейрона (до применения активационной функции), как:

$$y = \sum x_i w_i + b. \quad (1)$$

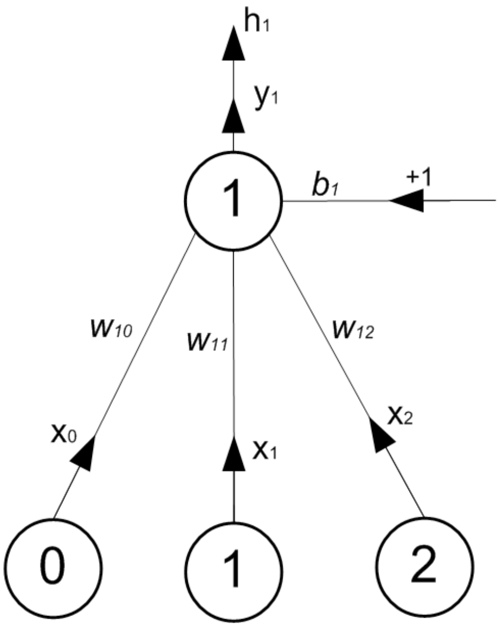


Рис. 1. Модель идеального нейрона, используемая в современных нейронных сетях глубокого обучения

Фундаментальным требованием к решению достаточно сложных задач распознавания с помощью искусственных нейронных сетей (построению нелинейных разделяющих поверхностей в пространстве признаков) является использование нелинейных активационных функции. Биологически обоснованным является использование сигмоидальной функции в качестве $f(y)$:

$$f(y) = \frac{1}{1+e^{-y}} . \qquad (2)$$

Однако на настоящий момент используются другие активационные функции[1]. Некоторые из них будут рассмотрены в следующем разделе.

[1] Это связано с эффектом «насыщения» сигмоидальной функции при значительном отличии сигнала x от 0 — значение градиента становится очень близким к нулю и во время обучения сигнал перестает распространяться через этот нейрон.

Таким образом, идеальный нейрон, подобно биологическому, имеет связи от выходов нескольких других нейронов (предыдущего слоя), осуществляет обработку и передает сигналы далее в сеть (в следующий слой). При этом следует отметить, что рассмотренная модель не функционирует во времени — отклик нейрона мгновенен, в отличие от реальных нейронов головного мозга. Протяженные во времени отклики моделируют с помощью спайковых сетей, однако практические результаты при применении этой модели весьма скромны, поэтому эти архитектуры далее подробно не рассматриваются.

Значения весов связей w_i и b получают в результате адаптации к набору изображений, для которых (в случае обучения с учителем) известен правильный ответ. Такой набор изображений называют обучающей выборкой, а набор изображений для контроля обучения — валидационной выборкой.

Обучение современных нейронных сетей в большинстве случаев осуществляется с помощью алгоритма обратного распространения ошибки [Rumelhart et al., 1985], который позволяет рассчитывать частные производные целевой функции (функции ошибки) по каждому весу w_i и b в нейронной сети. Зная эти производные, можно использовать алгоритм градиентного (наискорейшего) спуска для поиска локального минимума целевой функции в пространстве параметров. Рассмотрим алгоритм обучения подробнее.

Пусть E — функция потерь (ошибки). Для использования алгоритма наискорейшего спуска необходимо рассчитать частные производные по каждому свободному параметру $\partial E / \partial w_{ij}^k$, где w_{ij}^k — вес связи между нейроном j на слое $k-1$ с нейроном i на слое k. Для простоты пусть нелинейные активационные функции не применяются $h_i = y_i$. Алгоритм обратного распространения ошибки использует два прохода. Во время прямого прохода на вход НС подается изображение из обучающей выборки, формируются активации нейронов $\{h_i^k \in \mathrm{H}^k\}$, $k = [1, n]$, где n — количество слоев в нейронной сети. На выходе формируется вектор $\mathrm{H}^n = \{h_i^n\}$, который представляет собой «ответ» сети (например, метку класса) на пример из обучающей выборки, для которого заранее известно правильное реше-

ние $H^* = \{h_i^*\}$. Функция ошибки E является функцией от H^n и H^*, для которой можно аналитически рассчитать производную[2] $\partial E / \partial h_i^n = f(h_i^n, h_i^*)$.

При подстановке конкретных значений h_i^n и h_i^* это позволяет получить значение для $\partial E / \partial h_i^n$, которое можно использовать для расчета весов $\partial E / \partial w_{ij}^{n-1}$. Поскольку в нашем случае (без нелинейных акт. функций)[3] $h_i^n = \sum h_j^{n-1} w_{ij}^{n-1}$, а все значения h_i^k известны в результате прямого прохода, то, используя правило взятия сложной производной, можно рассчитать искомые значения для веса w_{ij}^{n-1}:

$$\frac{\partial E}{\partial w_{ij}^{n-1}} = \frac{\partial E}{\partial h_i^n} \frac{\partial h_i^n}{\partial w_{ij}^{n-1}} = \frac{\partial E}{\partial h_i^n} \frac{\partial(\sum h_j^{n-1} w_{ij}^{n-1})}{\partial w_{ij}^{n-1}} = \\ = \frac{\partial E}{\partial h_i^n} h_j^{n-1} = f(h_i^n, h_i^*) \cdot h_j^{n-1} \quad . \tag{3}$$

Можно также рассчитать и значения $\partial E / \partial h_i^{n-1}$, необходимые для расчета $\partial E / \partial w_{ij}^{n-2}$:

$$\frac{\partial E}{\partial h_j^{n-1}} = \sum_i \frac{\partial E}{\partial h_i^n} \frac{\partial h_i^n}{\partial h_j^{n-1}} = \sum_i \frac{\partial E}{\partial h_i^n} w_{ij}^{n-1} . \tag{4}$$

Сигнал можно распространять обратно к первым слоям сети, получая все необходимые производные $\partial E / \partial w_{ij}^k$. Наличие нелинейных активационных функций принципиально не усложняет задачу, а лишь добавляет промежуточный расчет производной $\partial E / \partial y_i$, поскольку активационные функции выбираются дифференцируемыми:

---

[2] В общем случае значение ошибки для конкретного выхода h_i может зависеть от всего вектора H и/или H^*.

[3] Можно представить вес b, как связь с нейроном, у которого активация постоянна и равна 1.

$$\frac{\partial E}{\partial y_i} = \frac{\partial E}{\partial h_i}\frac{\partial h_i}{\partial y_i} = \frac{\partial E}{\partial h_i}\frac{\partial f(y_i)}{\partial y_i} = \frac{\partial E}{\partial h_i} g(y_i)\,. \quad (5)$$

Для изменения весов в процессе обучения используют градиентный спуск с характерными адаптациями для улучшения скорости и качества сходимости. Рассмотрим один из популярных вариантов– компромиссный градиентный спуск с инерционным шагом обучения и регуляризацией размеров весов. На каждой итерации компромиссного градиентного спуска для расчёта направления градиента используется несколько изображений (градиент усредняется для каждого из изображений). Это позволяет выбирать направление спуска более точно, чем по одному примеру, при этом сохраняя высокую скорость обучения. На k-й итерации веса нейронной сети изменяются согласно следующей формуле:

$$w_{ik} = w_{ik-1} + \Delta w_{ik}\,, \quad (6)$$

где w_{ik} — i-ый вес нейронной сети на шаге k, а Δw_{ik} — изменение веса w_{i} на шаге k.

Инерционный шаг обучения (т.н. momentum) позволяет преодолевать локальные минимумы и повышает скорость. При использовании такого подхода веса на каждом шаге обучения меняются по следующему закону:

$$\Delta w_{ik} = \lambda \Delta w_{ik-1} - L_k \frac{\partial E}{\partial w_i}, \quad (7)$$

где λ — настраиваемый коэффициент, $\partial E / \partial w_i$ — производная функции ошибки по весу нейронной сети w_i, посчитанная с помощью алгоритма обратного распространения ошибки, L_k — шаг градиентного спуска (меняется по определенному закону так, чтобы со временем он уменьшался).

Популярен подход ограничения величин весов. В этом случае полагается, что целевая функция C состоит из функции ошибки E и дополнительного слагаемого, препятствующего чрезмерному увеличению весов:

$$C = E + \alpha / 2 \cdot \sum w_i^2\,, \quad (8)$$

Где w_i — это i-ый вес нейронной сети, а α– коэффициент регуляризации.

В этом случае:

$$\frac{\partial C}{\partial w_i} = \frac{\partial E}{\partial w_i} + \alpha w_i\,. \tag{9}$$

Таким образом, формула (7) с учетом новой целевой функции преобразуется в формулу (10).

$$\Delta W_k = \lambda \Delta w_{ik-1} - L_k \frac{\partial E}{\partial w_i} + L_k \alpha w_{ik-1}. \tag{10}$$

Указанные модель идеального нейрона и способ обучения нейронных сетей применимы для всех архитектур НС, рассмотренных далее для решения различных задач автоматического анализа изображений.

Нейронные сети для распознавания изображений

Глубокие архитектуры, основанные на искусственных нейронных сетях, берут свое начало с неокогнитрона, разработанного Кунихико Фукусимой в 1980 году [Fukushima, 1980]. В частности, можно отметить четыре важных особенности, положенные Фукусимой в основу его нейронных сетей глубокого обучения, которые были заимствованы в живых зрительных системах. Это [Луцив, 2015]:

- Использование детекторов специфических деталей изображения, реализованных на основе специальной формы рецептивных полей нейронов;
- Способ адаптивного обучения таких детекторов на основе предъявляемых нейронной сети характерных примеров зрительных образов;
- Реализация сверточных нейронных слоев на основе взаимодействия «простых» и «сложных» нервных клеток (именно в таких слоях выполняется вычисление свертки анализируемого зрительного образа с функцией, реализуемой нейронным детектором, или, иными словами, осуществляется корреляционный поиск характерных представляющих интерес деталей в анализируемой области изображения);

- Иерархическая организация процесса классификации изображения в сети глубокого обучения, в рамках которой на более высоких иерархических уровнях осуществляется обобщение информации, полученной от нескольких нейронных детекторов, функционирующих на более низких (предыдущих) иерархических уровнях.

К решению практических задач распознавания изображений первые искусственные нейронные сети были применены еще в 1989 году, когда Ян ле Кун с соавторами разработал сверточную нейронную сеть Lenet для распознавания рукописных цифр [LeCun et al., 1989, 1998]. Это решение впоследствии имело большой успех и было внедрено для автоматического распознавания отсканированных чеков. Основными элементами сверточной нейронной сети Lenet были: сверточные слои (СС), полносвязные слои (ПСС) и операции подвыборки.

В полносвязных слоях каждый нейрон текущего слоя связан с каждым нейроном предыдущего слоя. Вход y_i i-го нейрона вычисляется по формуле:

$$y_i = \sum_{j=1}^{N} x_j \cdot w_{ij} + b \tag{11}$$

где x_j — это выход нейрона предыдущего слоя, N — количество нейронов предыдущего слоя, w_{ij} – веса связей полученные в результате обучения, b — вес связи с постоянным значением. На настоящий момент полносвязные слои редко встречаются в современных архитектурах нейронных сетях глубокого обучения из-за сложности настройки большого количества весов в них содержащихся. Сверточные слои являются частным случаем полносвязных слоев и позволяют преодолевать указанный недостаток.

В сверточном слое каждый нейрон текущего слоя связан с ограниченным числом пространственно близких нейронов предыдущего слоя или пикселей изображения. Таким образом подобно нейронам сетчатки и зрительной коры формируются рецептивные поля идеальных нейронов в сверточных нейронных сетях. Важными по-

нятиями являются «фильтры» (ядра свертки) и «карты признаков» (feature maps). Свертка (convolution, conv) с k-м фильтром в сверточных слоях осуществляется по формуле:

$$y_{r,c,k} = \sum_{i=-H/2}^{H/2} \sum_{j=-W/2}^{W/2} \sum_{d=1}^{D} x_{i+r,j+c,d} \cdot w_{i,j,d,k} + b_k, \qquad (12)$$

где $(r;\ c)$ — это положение фильтра на изображении (или на «карте признаков» предыдущего слоя), $y_{r,c,k}$ — активация нейрона, соответствующего определенному положению фильтра k, x — выход нейронов предыдущего слоя, $H,\ W,\ D$– это высота, ширина и глубина фильтра (глубина фильтра, как правило, равна количеству карт признаков на текущем слое или количеству каналов изображения), $w_{i,j,d,k}$ — это веса фильтра k, полученные в результате обучения, b_k — вес связи с постоянным значением (связь с единицей) (также получается в результате обучения).

Совокупность $y_{r,c}$ (после применения активационной функции) для всех r и с (т.е. для всех положений фильтра) называют картой признаков, поскольку в результате обучения ядро свертки строится таким образом, чтобы реагировать на определенный признак изображения, например, вертикальную границу между темной и светлой областями изображения. Таким образом, карта признаков — это двумерная матрица откликов нейронов для определенного фильтра. Совокупность карт признаков определенного слоя формирует «стопку» карт, обрабатываемую дальше.

Важно отметить, что несмотря на то, что каждый нейрон имеет собственное рецептивное поле, реакция разных нейронов одной карты признаков на один и тот же стимул идентична, а поэтому во время обучения соответствующие веса разных нейронов в рамках одной карты должны оставаться одинаковыми. Это достигается за счет модификации алгоритма обратного распространения ошибки — на каждом шаге градиентного спуска для изменения веса $w_{i,j,d,k}$ используется сумма производных, рассчитанных для каждого из N нейронов соответствующей карты признаков:

$$\frac{\partial E}{\partial w_{i,j,d,k}} = \frac{1}{N} \sum_{r,c} \frac{\partial E}{\partial w_{i,j,d,k}(r,c)} . \tag{13}$$

Это позволяет значительно снизить количество настраиваемых параметров, а, следовательно, уменьшить вероятность переобучения нейронной сети.

Пулинг (pooling) процедура — это способ осуществления подвыборки данных, которая позволяет уменьшить пространственное разрешение обрабатываемого сигнала, а также сделать работу сети нечувствительной к незначительным локальным сдвигам изображения. Основные операции такого рода:

Max pooling — по каждой из карт признаков с определенным шагом проходит скользящее окно определенного размера. На следующий слой передаются только максимальные значения карты признаков внутри окна.

Average pooling — аналогично, но передается далее среднее значение в окне.

В 2012 году в конкурсе по распознаванию базы данных (БД) ImageNet [ImageNet Large Scale Visual Recognition Challenge, 2012] была впервые успешно применена искусственная нейронная сеть. ImageNet содержит более миллиона изображений тысячи различных классов объектов, среди которых животные, растения, техника, здания и предметы мебели. Решение на основе архитектуры [Krizhevsky et al., 2012] победило с большим отрывом. Примеры изображений из этой базы данных приведены на рис. 2.



Рис. 2. Пример изображений ImageNet

Решение по сути сохраняло все основные черты Lenet, но сама сеть была намного больше. Описание архитектуры приведено на рис. 3.

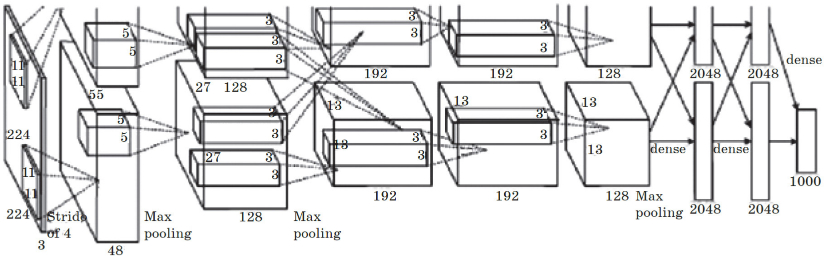


Рис. 3. Архитектура сверточной нейронной сети, победившей в соревновании ImageNet 2012 года

Сеть [Krizhevsky et al., 2012] имеет семь скрытых слоев (пять сверточных и два полносвязных). Фильтры для обнаружения характерных локальных элементов изображения формируются автоматически в процессе обучения сети на представительной выборке изображений объектов большого количества классов и адаптированы к свойствам изображений использованной обучающей выборки. Как можно легко заметить по рис. 4, фильтры-детекторы первого слоя (работающие непосредственно с изображением), напоминают вейвлеты Габора или маски Лавса [Laws, 1980].

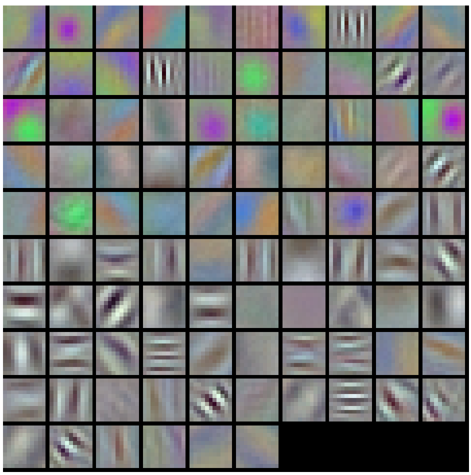


Рис. 4. Фильтры детекторы, построенные нейронной сетью в результате ее обучения

Тем не менее, одного скрытого слоя было бы явно недостаточно для обеспечения таких прекрасных результатов. Наибольший интерес для анализа могут представлять фильтры, сформированные в более глубоких скрытых слоях сети (начиная со второго), однако их интерпретация намного сложней. В 2013 году исследование по «пониманию» нейронных сетей глубокого обучения было выполнено с использованием «инверсной свертки» [Zeiler et al., 2010]. Используя сеть с «инверсной сверткой», авторы [Zeiler, Fergus, 2014] проецируют значения, вычисленные в картах признаков различных сверточных слоев, обратно в соответствующие части слоя входного изображения. Таким образом, авторы определили, за анализ каких фрагментов входного изображения «отвечают» конкретные нейроны конкретной карты признаков и как количество и свойства нейронов сверточных слоев влияют на адекватность проецирования входного изображения в конкретную карту признаков. На основе такого анализа были оптимизированы параметры сверточной сети, за счет чего такое техническое решение победило в соревновании ImageNet-2013. Важным стало то, что при исследовании сетей с «инверсной сверткой» их создатели выяснили, что «образы» проекций активности выходных нейронов начальных сверточных слоев на слой входного изображения соответствуют образам промежуточных представлений локальных элементов изображения в живых зрительных системах, как это проиллюстрировано на рис. 5 и описано в монографии Д. Марра [Марр, 1987].

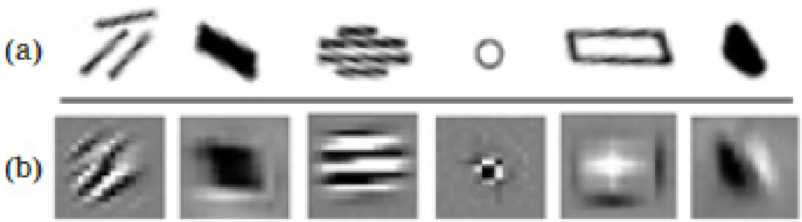


Рис. 5. (*а*) — Представления локальных элементов изображения в живых зрительных системах, описанные Марром, (*б*) — «образы», полученные обратным проецированием активности первых сверточных слоев на слой исходного изображения

Таким образом, признаки, автоматически сформированные в первых слоях нейронной сети в результате автоматического обучения, близки к естественным и наилучшим теоретическим решениям. Можно (хоть это и не очевидно) предположить, что и фильтры, сформированные в более высоких слоях отвечающие за структурный анализ изображения, также являются в некотором роде оптимальными, как показали исследования [Zeiler et al., 2010] они уже интерпретируются в рамках частей объектов и объектов на изображении.

О высоких обобщающих способностях выученных признаков больших нейронных сетей говорит хорошая переносимость признаков на другие задачи. Так, например, в той же работе [Zeiler et al., 2010] был получен лучший на то время результат по распознаванию классов некоторых других баз данных путем «перенесения» признаков сети, обученной распознаванию тысячи классов ImageNet.

В [Krizhevsky et al., 2012] было показано, что ректифицированная линейная активационная функция RELU (rectified linear unit) во многих случаях позволяет сети обучаться более эффективно, чем сигмоида. Эта активационная функция вычисляется по формуле:

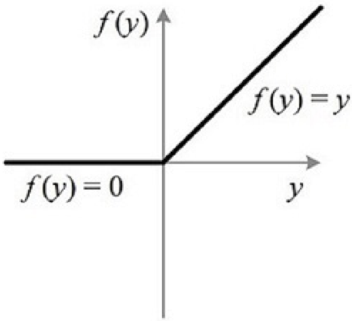
$$f(y) = max(0, y)\,. \tag{14}$$

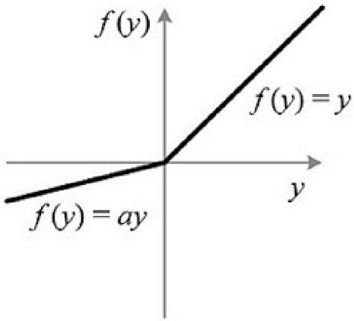
Еще более современная активационная функция — PRELU, — трактует наклон отрицательной части функции, как свободный параметр в процессе обучения алгоритмом обратного распространения ошибки [He et al., 2015]. PRELU вычисляет значение отклика нейрона, согласно следующей формуле:

$$f(y_i) = \begin{cases} y_i, y_i > 0 \\ \alpha_i y_i, y_i \le 0 \end{cases}, \tag{15}$$

где y_i — это отклик нейрона в i-й карте признаков, а α_i — коэффициент, управляющий наклоном отрицательной части, обучаемый как свободный параметр во время выполнения алгоритма обратного распространения ошибки.

Эксперименты [Malashin, Anisimov, 2016] показали, что использование PRELU вместо RELU (рис. 6) позволяет ощутимо увеличить точность распознавания валидационной выборки, а также немного превосходит другие модификации RELU — биноминальную нор-





(*a*) (*b*)

Рис. 6 Активационные функции, используемые в современных искусственных нейронных сетях: (*а*) RELU, (*б*) PRELU

мальную логистическую активационную функцию правдоподобия (binominal normal log likelyhood — BNLL), и активационную функцию, аналогичную PRELU, но с фиксированным параметром α_i (Leaky RELU).

Общее количество настраиваемых параметров в ходе обучения сети [Krizhevsky et al., 2012] превысило 150 000 000. Для возможности обучения таких больших искусственных нейронных сетях применяют параллельные вычисления на графическом ускорителе. На настоящий момент лучшие архитектурные решения, как правило, требуют меньшего количества весов, но являются гораздо более глубокими. Это достигается, в том числе, за счет применения сверток 1×1, т.е. фильтров с шириной и высотой равной единице. Было замечено [Lin et al., 2013], что классические сверточные слои вычисляют отклик как линейную комбинацию активаций предыдущего слоя, а преобразования, которым подвергаются локальные участки изображения, являются нелинейными. Использование многослойных перцептронов вместо ядер свертки делает возможным отображение фрагментов изображения и карт признаков на нелинейное многообразие, что в случае классических сверточных слоев возможно лишь с помощью нескольких слоев. При этом реализация таких архитектур возможна с помощью комбинации сверток N×N со свертками 1×1. Нейронные сети с такими слоями получили название «Сеть

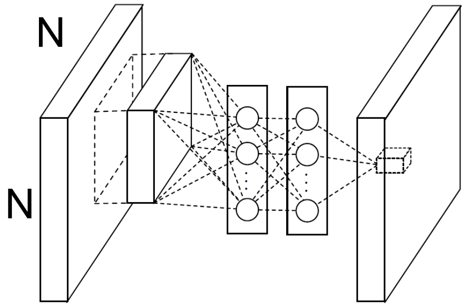


Рис. 7. Архитектура сверточного слоя сеть в сети размером N×N

в Сети» (Network in Network) (рис. 7).

Первая работа, в которой акцент делался на увеличении глубины сети, стала GoogleNet [Szegedy et al., 2014], выполненная в компании Google — количество слоев было увеличено до 24 (по сравнению с 7 в сети [Krizhevsky et al., 2012]), при этом в одном слое комбинировались свертки с ядрами разных размеров. Разработчики назвали эту архитектуру «сеть в сети в сети» («network in network in network»). Это решение победило на соревновании ImageNet в 2014 году. Из-за большой глубины при обратном распространении ошибки сигнал «размывается» к первым слоям, что делает обучение не эффективным. Для преодоления этого эффекта в GoogleNet распознавание происходит не только на верхнем слое, но также и на классификаторах — нескольких надстроек к промежуточным слоям сети. Таким образом подразумевается, что промежуточные слои также должны содержать достаточно дискриминантные признаки, по которым можно принимать решение.

Альтернативой построению нескольких «классификаторов» внутри одной сети является добавление прямых связей, передающих сигнал без изменений в обход фильтров свертки. На таком принципе основана сеть-победитель соревнований ImageNet-2015 ResNet [He et al., 2015], которая содержит слои «пропуска», передающие активацию в обход сверточных фильтров и суммирующие ее с результатом свертки.

Другими словами, ResNet формируют активацию x_l на слое l, как:

$$x_l = H_l(x_{l-1}) + x_{l-1} \quad (16)$$

где H_l — составная функция, которая в конкретном случае, представляет собой последовательное использование свертки с определенным ядром и нелинейной активационной функции.

Глубина ResNet составляет более 150 слоев, глубина ещё более новой архитектуры DenseNet [Huang et al., 2017] для распознавания ImageNet может достигать 201 слоя. «Плотносвязные блоки» DenseNet состоят из последовательного комбинирования пары чередующихся сверточных фильтров размером 1×1 и 3×3, при этом карты признаков, полученные после каждой пары, дублируются, как входные для всех последующих.

Для DenseNet активация на слое lвыражается как:

$$x_l = H_l([x_0, x_1, ..., x_{l-1}]), \quad (17)$$

где H_l представляет собой нормировку [Ioffe, Szegedy, 2015], линейную ректифицированную функцию и свертку с ядром 3×3. При этом в H перед каждым ядром свертки 3×3 также используется свертка 1×1 с меньшим количеством фильтров, чем входных карт признаков. Это позволяет ограничить экспоненциальный рост количества карт признаков от слоя к слою.

Между плотносвязными блоками используются так называемые слои перехода, которые с помощью фильтров 1×1 и операции max pooling понижают разрешение и ширину «стопки» карт признаков между блоками. Пример архитектуры сети DenseNet с тремя плотносвязными блоками приведен на рис. 8.

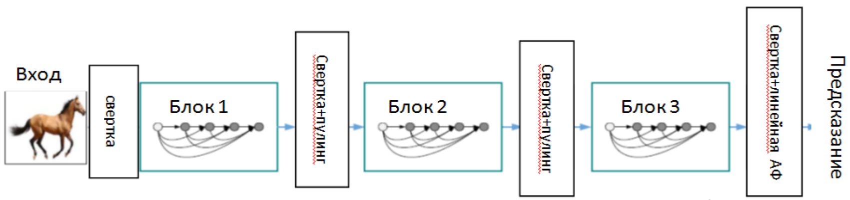
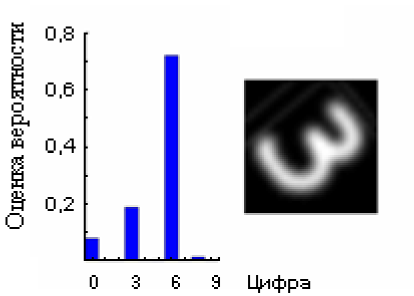


Рис. 8. Нейронная сеть DenseNet с тремя блоками

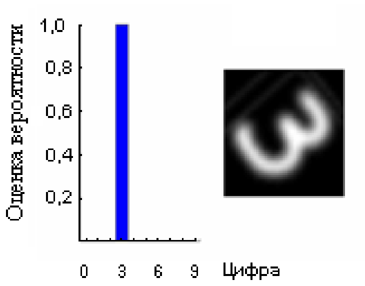
Точность распознавания базы данных ImageNet (как и многих других баз данных меньшего размера) постоянно увеличивается за счет улучшения архитектур нейронных сетей глубокого обучения. На настоящий принято уже говорить о «сверхчеловеческой» точности распознавания достигнутой в этом конкурсе (например, [He et al., 2015]) — ошибка распознавания валидационной выборки (при выборе пяти наиболее вероятных классов из 1000) составляет менее 5%, что меньше чем в эксперименте с человеком[4].

Тем не менее, при использовании нейронных сетей на практике следует иметь ввиду, что очень большое значение имеет использованная обучающая выборка, поскольку обобщающие способности нейронных сетей глубокого обучения ограничены. Такая ограниченность (при обобщении геометрических искажений) была описана в работе [Малашин, Кадыков, 2015]. Согласно полученным, выводам, для достижения инвариантного к повороту распознавания необходимо, чтобы в обучающей выборке находились изображения во всем диапазоне возможных поворотов. Очевидным объяснением такого вывода является тот факт, что в условиях априорной неопределенности каких-либо параметров процесс распознавания или обнаружения является более сложным сам по себе. Это в полной мере справедливо как для алгоритмов автоматического анализа, так и для человека [Красильников, 1986]. Нейрофизиологические исследования подтверждают справедливость такого заключения и для распознавания образов в гауссовом шуме [Байдаков и др., 1973]. Таким образом, предположение о неопределенности или дополнительной изменчивости какого-либо из параметров изображения приведет к обязательному ухудшению распознавания обучающей выборки, и, следовательно, это допущение становится нецелесообразным согласно классически используемой целевой функции, оценивающей качество распознавания изображения только на обучающей выборки. Пример, демонстрирующий распознавание в условиях априорной неопределенности ориентации изображения на исходной и подготовленной обучающей выборке, представлен на рис. 9.

[4] Характеристика «сверхчеловеческое качество распознавания» в этом случае вряд ли является объективной, хотя бы потому, что эталонные метки приписываются изображению человеком, а точность распознавания для разных людей разная.



(*а*)



(*б*)

Рис. 9. Распознавание в условиях априорной неопределенности ориентации изображения с помощью сверточной нейронной сети глубокого обучения. Для эксперимента использовалась база данных рукописных цифр MNIST. (*а*) — тестовое изображение и результат при обучении на исходной обучающей выборке. (*б*) — при расширении обучающей выборки повернутыми изображениями.

Таким образом, преобразование изображений, извлекаемых из разнородных источников, необходимо для того, чтобы обеспечивать максимально возможную схожесть обучающей выборки с изображениями, полученными при использовании нейронной сети на практике.

Зачастую этот факт не учитывается разработчиками. Так, например, в статье [Hosseini et al., 2017] показано, что недавно представленный программный интерфейс компьютерного зрения от компании Google[5] не устойчив к аддитивному импульсному шуму — подверженные такому шуму изображения из базы данных ImageNet перестают распознаваться (рис. 10).

В качестве решения авторы [Hosseini et al., 2017] предлагают использование внешнего шумоподавляющего алгоритма, после применения которого изображения вновь начинают классифицироваться корректно. Однако можно заметить, что нейронные сети сами способны выучить алгоритм шумоподавления (что успешно делают с помощью шумоподавляющих автоэнкодеров[6]), а описанный эф-

[5] С учетом научного состояния вопроса закономерно предположить, что ядром используемого классификатора являются сверточные нейронные сети.

[6] Подробнее об автоэнкодерах рассказывается в разделе, посвященном удалению шума с изображений.

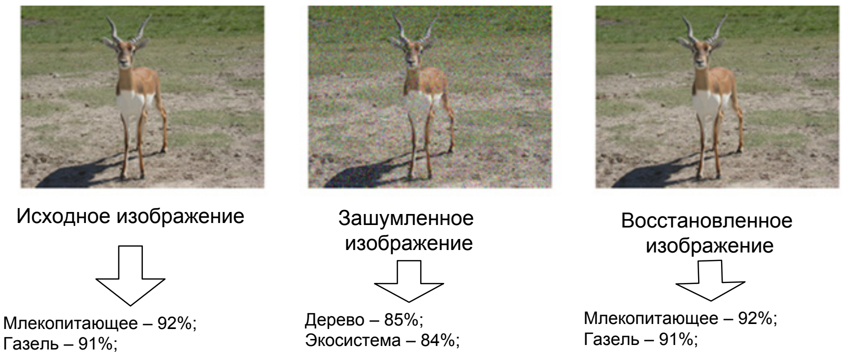


Рис. 10. Демонстрация неустойчивости алгоритмов Google к шуму на изображении [Hosseini et al., 2017]. Справа исходное изображения и метки, приписанные ей алгоритмом (млекопитающее, газель), по центру изображение, подверженное шуму (дерево, экосистема), справа — результаты распознавания после применения внешнего шумоподавляющего алгоритма (млекопитающее, газель)

фект можно трактовать также, как эффект переобучения — появление шума воспринимается как признак определенного класса. Об этом, в частности, свидетельствует то, что в [Hosseini et al., 2017] для двух семантически разных изображений при добавлении шума была приписана одна и та же метка.

Нейронные сети для обнаружения объектов на изображении

Поскольку ImageNet изначально была предназначена для обучения классификации изображений (приписывания всему изображению целочисленной метки, соответствующей классу представленного на нем объекта), то и основные успехи применения искусственных нейронных сетей были связаны с этой задачей. В связи с этим, создание успешных алгоритмов обнаружения (приписыванию изображению меток объектов и соответствующих им описывающих прямоугольников[7]) с использованием нейронных сетей первоначально осуществляли с помощью построения специальной «надстройки» над алгоритмом классификации. До настоящего времени лишь небольшое количество решений использует обучение глубо-

[7] Точность приписанного описывающего прямоугольника принято оценивать по метрике $IoU(a,b)$ (IntersectionoverUnion) – отношению площадей пересечения эталонного (a) и приписываемого (b) прямоугольника к их объединению. $0 \leq IoU(a,b) \leq 1$.

кой нейронной сети-детектора без так называемой базовой (опорной) нейронной сети — нейронной сети, обученной классификации на базе данных ImageNet. В качестве опорной нейронной сети, как правило, используют архитектуры VGG [Simonyan, Zisserman, 2014], GoogleNet [Szegedy et al., 2014] или ResNet [He et al., 2015].

Из-за сравнительно небольших размеров баз данных с информацией о положении объектов для обучения нейронных сетей-детекторов широко используют стратегию расширения обучающей выборки с помощью, вращения, изменения цвета и яркости (в заданных диапазонах). Другой широко используемой техникой является поиск сложных ложно-позитивных примеров.

Для задачи обнаружения широкое применение нашла региональная сверточная нейронная сеть R–CNN (Region based Convolutional Neural Network) [Girshick et al., 2015]. Архитектурное решение было основано на использовании внешнего алгоритма сегментации, результатом которого являлся сгенерированный набор фрагментов исходного изображения, которые (предварительно отмасштабировав к фиксированному размеру) можно подавать на вход сети, обученной классификации изображений ImageNet. При превышении определенного порога вероятности для одного из классов можно считать, что объект обнаружен в прямоугольнике, соответствующем представленному фрагменту. Для ускорения процесса, а также лучшей адаптации к классификации отдельных фрагментов, опорную нейронную сеть доучивали на сравнительно небольшой базе данных Pascal [The PASCAL Visual Object Classes], содержащей эталонные описывающие прямоугольники. При этом окончательную классификацию осуществляли на основе признаков верхнего сверточного слоя с помощью SVM.

Предложенный в [Girshick et al., 2015] подход позволял полноценно использовать впечатляющие результаты классификации для решения задачи обнаружения, однако имел ряд недостатков:

1. В этом случае требуется многократный запуск нейронной сети на одном изображении, что ведет к неоправданно большим затратам вычислительных ресурсов.
2. Нейронная сеть лишается возможности учитывать контекст (область вокруг) объекта интереса.

3. Узким местом становится алгоритм сегментации, который не учитывает семантическую информацию[8].

Некоторых улучшений скорости удалось добиться с помощью подхода, получившего название Fast R-CNN (быстрая региональная сверточная нейронная сеть) [Girshick, 2015]. В этом случае при обучении на вход опорной нейронной сети подается все изображение целиком (а не фрагмент), после этого рассматривается карта признаков одного из верхних слоев, из которых в соответствии с предлагаемым внешним алгоритмом регионом интереса произвольного размера извлекают вектор признаков фиксированного размера. Последнее действие не является тривиальным. В предложенном автором [Girshick, 2015] способе оно основано на использовании специального слоя «пулинга областей», который делит произвольную по размеру карту признаков сеткой с заданным количеством ячеек и передает дальше максимальный отклик нейрона в каждой ячейке. Далее отклики преобразуются последовательно с помощью нескольких полносвязных слоев в вектор признаков, который в свою очередь передается в две локально связные ветки, в результате чего получается два «родственных» выхода:

1. Распределение вероятностей классов по K+1 категории $p = (p_0, \ldots, p_K)$.
2. Вектор из K четырехкомпонентных наборов, в которых закодировано инвариантное к масштабу смещение, и логарифм ширины и высоты описывающего прямоугольника

$$t^k = (t_x{}^k, t_y{}^k, t_w{}^k, t_h{}^k)\,.$$

Общая схема архитектуры приведена на рис. 11.

При обучении целевая функция для эталонной метки u и эталонного описывающего прямоугольника (ОП) v имеет вид:

$$L\left(p,u,t^u,v\right) = L_{cls}\left(p,u\right) + \lambda f\left(u\right) L_{loc}\left(t^u,v\right), \tag{18}$$

где L_{cls} — ошибка классификации, L_{loc} — ошибка локализации, t^u — описывающий прямоугольник для истинного класса u, а функ-

[8] Несомненно, что эта информация очень важна для правильной сегментации. Об этом, в частности, свидетельствуют успехи нейронных сетей глубокого обучения при решении этой задачи (см. следующий подраздел).

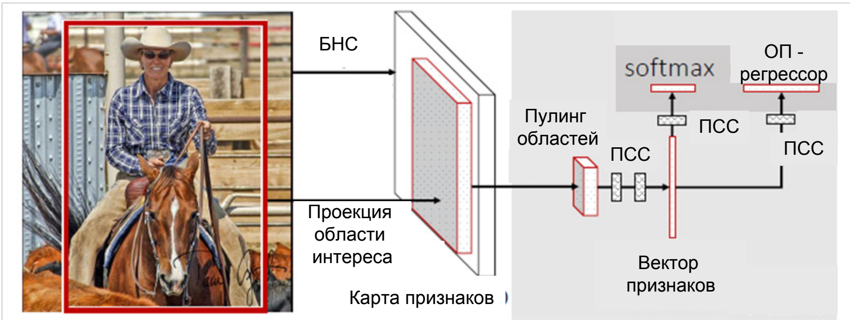


Рис. 11. Архитектура нейронной сети Fast R-CNN

ция $f(u)=1$, для всех нефоновых классов и принимает значение 0 в противном случае. $L_{cls}(p,u)=-\log(p_u)$ для истинного класса u.

$$L_{loc}\left(t^u,v\right)=\sum_{i\in\{x,y,w,h\}} smooth_{L_1}(t_i^u-v_i), \tag{19}$$

$$smooth_{L_1}(x)=\begin{cases}0.5x^2, \text{ если } |x|<1\\ |x|-0.5, \text{ иначе}\end{cases}. \tag{20}$$

Следующая модификация R–CNN — Faster R–CNN [Ren et al., 2016]. Она использует для генерации потенциальных областей интереса вместо селективного поиска — алгоритма, не учитывающего семантическое содержание изображения — другую нейронную сеть, предлагающую несколько областей интереса для каждого положения на рассматриваемой карте признаков (Region Proposal Network (RPN)). По сути RPN представляет собой небольшую нейронную сеть, состоящую из нескольких полносвязных слоев, принимающую на вход фрагмент карты признаков опорной нейронной сети размером n×n и выдающую вектор вероятностей и набор закодированных координат описывающих прямоугольников. По сути RPN — это детектор в скользящем окне — в каждом положении фрагменты карты признаков обрабатывается одинаково (веса нейронной сети фиксированы для всех положений RPN на карте признаков). Стоит отметить, что подобная архитектура может быть от начала и до

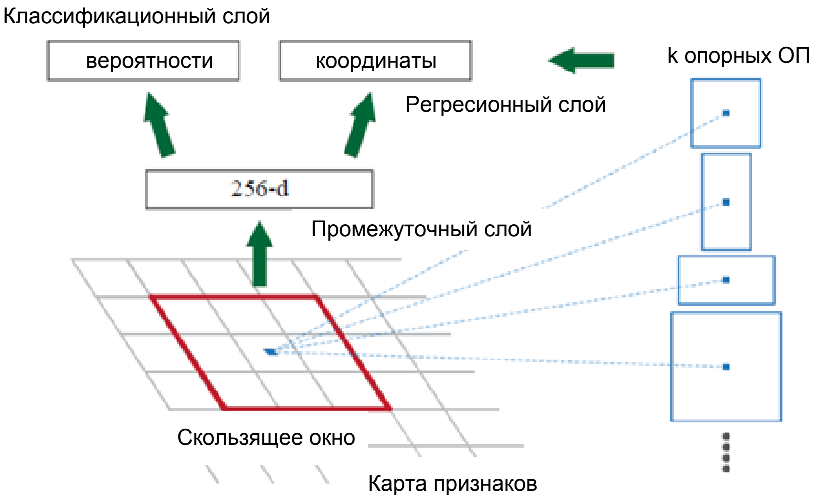


Рис. 12. Иллюстрация архитектуры RPN

конца эффективно реализована на стандартных библиотеках глубокого обучения, подобных Caffe [Caffe Deep Learning Library], за счет использования сверточных слоев размером n×n и 1×1 (хотя вопрос обучения такой сети из-за наличия слоя «пулинга областей» не однозначен [Ren et al., 2016]). Еще одним нововведением является использование опорных описывающих прямоугольников (ОП) — набора из нескольких описывающих прямоугольников (ОП) разного размера и соотношения сторон, описывающих положение объектов относительно центра скользящего окна «по умолчанию». Другими словами, они описывают некоторое априорное предположение о том, где может находится объект. RPN же предсказывает смещение относительно k опорных описывающих прямоугольников (ОП), а также вектор размером 2k, содержащий результат 2-классового предсказания («объект–не объект»). За счет использования опорных описывающих прямоугольников (ОП) разного размера отпадает необходимость обрабатывать изображение на разных масштабах. Этот подход получил широкое распространение и в других архитектурах. Авторы [Ren et al., 2016] используют девять опорных положений (три масштаба и три соотношения сторон) и таким образом общее

количество предсказываемых областей интереса для всей нейронной сети составляет W×H×9, где W×H — размер обрабатываемой карты признаков. Рис. 12 иллюстрирует работу RPN.

В процессе обучения целевая функция учитывает три вида опорных описывающих прямоугольников (ОП):

– если описывающий прямоугольник (ОП) содержит объект, он считается положительным (определяется по относительной площади пересечения и объединения описывающего прямоугольника (ОП) и эталонного описывающего прямоугольника (IoU > 0.7));
– если описывающий прямоугольник (ОП) не содержит объект, он считается отрицательным (IoU < 0.3);
– все прочие описывающие прямоугольники (ОП) не являются ни положительными, ни отрицательными.

Только положительные и отрицательные описывающие прямоугольники (ОП) вносят вклад в целевую функцию обучения. Каждый исходный описывающий прямоугольник параметризуется в вектор $t = \{t_x, t_y, t_w, t_h\}$ согласно правилу:

$$\begin{aligned} t_x &= (x - x_a) / w_a, \\ t_y &= (y - y_a) / h_a, \\ t_w &= \log(w / w_a), \\ t_h &= \log(h / h_a), \end{aligned} \tag{21}$$

где x, y, w, h — это координаты центра описывающего прямоугольника, ширина и высота; x, x_a — координата по оси абсцисс для предсказываемого и опорного описывающих прямоугольников (ОП) соответственно (аналогично для y, w, h).

Для каждого i-го отрицательного или положительного описывающего прямоугольника (ОП) ti и соответствующей ему предсказываемой вероятности p_i быть объектом функция ошибки L:

$$L\left(\{p_i\},\ \{t_i\}\right) = \frac{1}{N}\sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*), \tag{22}$$

$$L_{cls} = -p^* \log(p) - \left(1 - p^*\right)\log(1 - p), \tag{23}$$

$$L_{reg} = smooth_{L_1}(t_i - t_i^*), \qquad (24)$$

где t_i^* — эталонный вектор, p_i — предсказываемая вероятность описывающего прямоугольника (ОП) быть объектом, эталонная метка p^* равна 1, когда опорный описывающий прямоугольник (ОП) положительный, и $p^* = 0$ в противном случае.

С учетом возможности имплементации RPN в виде нескольких сверточных слоев, можно говорить о изменении парадигмы R–CNN в Faster R–CNN, поскольку выделение шага «детекции» является не таким явным. RPN — это по сути часть объединенной сети, которая обучается задаче регрессии изображения в набор описывающих прямоугольников и вероятностей (хотя использования «пулинга областей» вызывает некоторые сложности с обучением). Таким путем изначально развивалась архитектура нейронной сети YOLO (you look only once, ты смотришь только один раз) [Redmon et al., 2015]. Эта архитектура расширяет опорную нейронную сеть с помощью нескольких полносвязных слоев, а на выходе имеет матрицу размером 7×7×(2*5+C), где, C — количество классов. Все изображение делится сеткой на 7×7 ячеек и для каждой ячейки предсказывается два описывающих прямоугольника (ОП). При этом оптимизируется квадратичная ошибка как по координатам описывающего прямоугольника (ОП), так и по ошибке приписывания метки (это является минусом, поскольку известно, что для классификации лучше подходят другие метрики). Архитектура YOLO представлена на рис. 13.

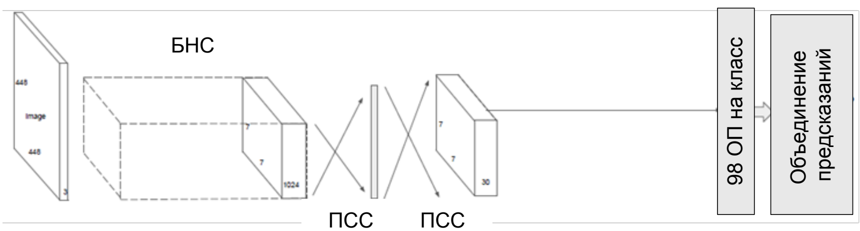


Рис. 13. Архитектура YOLO. (БНС — базовая (опорная) нейронная сеть, ПСС — полносвязный слой, ОП — описывающий прямоугольник)

В модификации YOLOv2 [Redmon, Farhadi, 2016] по сравнению с первой версией внесены изменения, которые делают эту архитектуру более схожей с Faster R–CNN — присутствуют только сверточные слои, используются опорные описывающие прямоугольники (ОП), однако в отличие от Faster R–CNN обнаружение и распознавание не разделены. При этом опорные описывающие прямоугольники (ОП) рассчитываются путем кластеризации положений и соотношений сторон объектов в обучающей выборке алгоритмом K-средних. YOLOv2, аналогично первой версии YOLO, делит изображение ортогональной сеткой, каждой ячейке которой соответствует «столбец» в карте признаков последнего сверточного слоя. От сети требуется, чтобы объект обнаруживали нейроны, которые соответствуют ячейке, находящейся по центру объекта и, таким образом, значения активаций верхнего слоя (после применения логистической функции σ) лежат в диапазоне от 0 до 1. Всего для описывающего прямоугольника (ОП) (ОП), предсказываемого сетью, формируется пять параметризованных координат $t = \{t_x, t_y, t_w, t_h\}$ и t_o. Если ячейка смещена на (c_x, c_y) и p_w и p_h — ширина и высота *опорного* описывающего прямоугольника (ОП), то:

$$b_x = \sigma(t_x) + c_x, \tag{25}$$

$$b_y = \sigma(t_y), \tag{26}$$

$$b_w = p_w e^{t_w} + c_y, \tag{27}$$

$$b_h = p_h e^{t_h}, \tag{28}$$

$$\Pr(object) \cdot IoU(b, object) = \sigma(t_o), \tag{29}$$

где $b = \{b_x, b_y, b_w, b_h\}$ — координаты описывающего прямоугольника в плоскости изображения, предсказываемые сетью. Пояснения к формулам даны на рис. 14.

Количество фильтров N предпоследнего слоя определяется количеством k опорных описывающих прямоугольников (ОП) и C классов: $N = k \cdot C + 5 \cdot k$. Использование полностью сверточной нейронной сети (т.е. сети, в которой используются только сверточные слои и слои pooling) делает активацию каждого слоя (вплоть до слоя

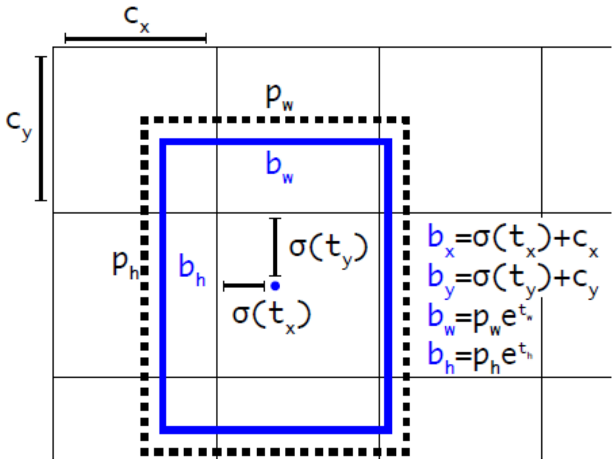


Рис. 14. Использование опорных описывающих прямоугольников в YOLOv2

с нейронами-детекторами) эквивариантной смещению изображения — сдвиг изображения ведет только к сдвигу активации на любом слое. Это позволяет масштабировать сеть для изображений разных размеров путем простого расширения каждой карты признаков на заданное количество столбцов. Этот факт используется YOLOv2 при обучении сети распознаванию объектов разных масштабов — сеть изменяет размер входного изображения каждые 10 итераций (на один из заранее определенных). Для того чтобы сеть могла принимать решение с учетом карты признаков с более высоким пространственным разрешением, в YOLOv2 добавляется также так называемый «пропускной» слой. Он переносит признаки с карты признаков предпоследнего слоя в неизменном виде на последний слой, по которому принимается решение. Это достигается за счет «вытягивания» карты — смежные четыре активации преобразуются в четыре разных канала карты признаков, после этого они стыкуются с картой признаков последнего признакового слоя.

Еще одним нововведением в YOLOv2 является использование иерархии объектов и комбинированного обучения классификатора и детектора — YOLO9000, который способен обнаруживать 9000 разных классов объектов. Для этого авторами было создано дерево классов, которое позволяет связать объекты нижнего уровня (например, «йор-

кширский терьер») с объектами более высокого уровня (например, «собака» или «животное»). Каждому объекту из обучающей выборки приписывается также все метки более высоких уровней, а на верхнем слое применяется функция Softmax, причем для каждого уровня независимо, что дает в общей сложности 9418 классов. Функция потерь, как и в других архитектурах, состоит из двух частей — ошибки классификации и ошибки обнаружения. Если в базе данных изображений не содержится информация о положении объекта на изображении, то вторая часть игнорируется. Это позволяет осуществлять комбинированное обучение сети одновременно на базе данных для классификации и детекции и, таким образом, обучаться обнаружению объектов, не имеющих эталонных описывающих прямоугольников (ОП). Другими словами, нейронной сети предоставляется возможность обобщать знания о схожести объектов при их обнаружении. По заявлению авторов, хорошее обобщение достигается на изображения животных, и плохое — для элементов одежды и снаряжения. Недостатком YOLO9000 является громоздкость, так, например, верхний сверточный слой, несмотря на уменьшение количества описывающих прямоугольников (ОП) до 3, содержит 28269 фильтров.

Основным преимуществом YOLOv2 является высокая скорость работы при сохранении достаточно высокой точности (рис. 15).

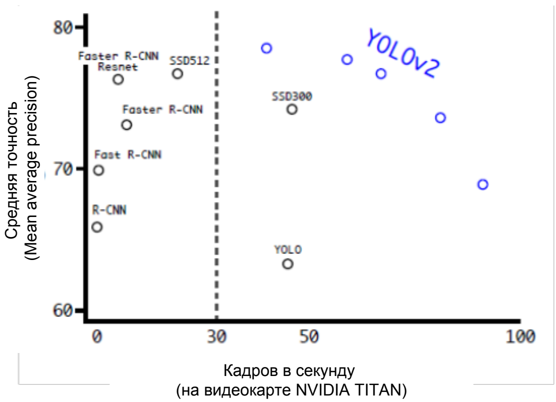


Рис. 15. Сравнение скорости работы разных детекторов и точности их работы на тестовой выборке VOC2007. Для YOLOv2 отмечено несколько значений, соответствующих разному разрешению обрабатываемого изображения

Еще одним детектором, осуществляющим регрессию изображения в набор координат описывающих прямоугольников, является SSD (Single Shot Detector) [Liu et al., 2015]. Его отличительной особенностью является использование нескольких карт признаков, надстроенных над опорной нейронной сетью, по которым принимается решение о нахождении объектов на изображении. Это делается с помощью детекторов-сверток (различных для разных карт признаков) размером 3×3. Архитектура SSD приведена на рис. 16.

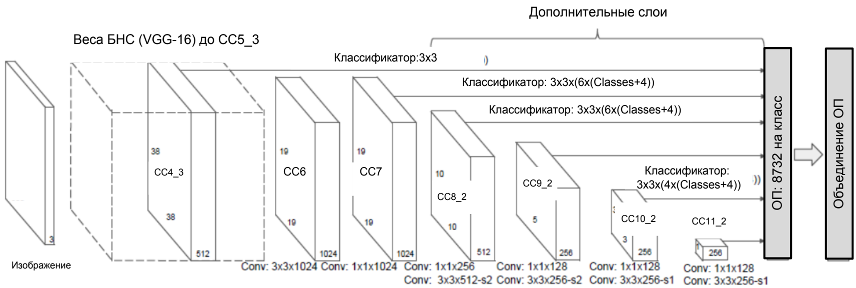


Рис. 16. Архитектура SSD. (СС (Conv) — сверточный слой, ОП — описывающие прямоугольники, БНС — базовая (опорная) нейронная сеть). В качестве опорной нейронной сети используется VGG-16, над которой надстраиваются несколько сверточных слоев. Решение принимается с помощью классификаторов-сверток по нескольким картам признаков добавленных слоев

Плюсом принятия решения по нескольким картам признаков является не только возможность учитывать признаки более высокого разрешения, но и эффективное (прямое) распространение градиента в процессе обучения обратным распространением ошибки, что значительно повышает его эффективность.

Основываясь на этой идее, DSOD (Deeply Supervised Object Detection) [Shen et al., 2017] расширяет архитектуру SSD с помощью подхода DenseNet [Huang et al., 2017]. Основой подобной архитектуры является передача некоторого количества активаций нейронов далее в сеть без изменений.

DSOD использует вариант DenseNet в качестве опорной сети, однако принципиально важным отличием от других сетей является то, что в этом случае опорная нейронная сеть не предобучается распоз-

наванию объектов, а учится вместе с надстройкой-детектором, которая представляет собой набор из четырех блоков [каждый блок — два сверточных фильтра (1×1 и 3×3)]. Карты признаков надстройки также передаются всем последующим слоям на вход. При этом поскольку происходит понижение разрешения (за счет перестановки сверточного фильтра с шагом 2), то на каждый последующий блок передается не исходная карта, а обработанная с помощью операции понижения размерности — max pooling и свертки 1×1[9].

Архитектура DSSD (Deconvolutional Single Shot Detector) расширяет SSD путем добавления деконволюционных слоев[10], интерполирующих карты признаков нижележащих слоев. В этом случае в качестве опорной нейронной сети используют архитектуру [Huang et al., 2017]. Деконволюционные модули имеют входы не только от карты признаков предыдущего слоя, но также и от обработанной особым образом карты признаков с тем же разрешением, который требуется восстановить. Это также способствует лучшему распространению градиента при обучении. Архитектура сети DSSD представлена на рис. 17.

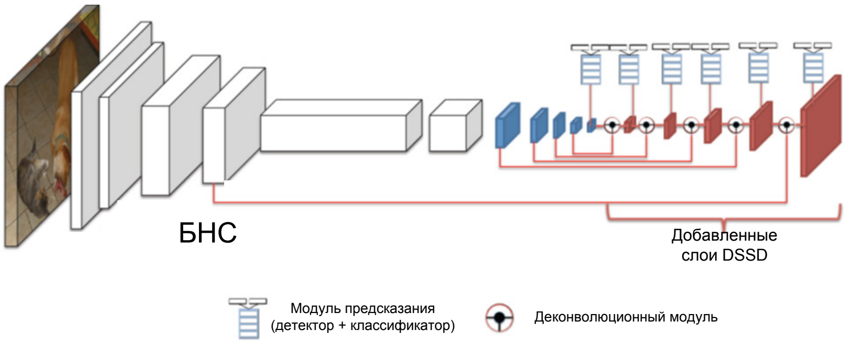


Рис. 17. Структура сети DSSD. Синим обозначены сверточные слои SSD, красным — соответствующие им деконволюционные признаковые слои. (БНС – базовая (опорная) нейронная сеть)

[9] Свертка 1×1 в данном случае выполняется с меньшим количеством фильтров, чем исходная карта признаков, и тем самым получается уменьшить ширину «стопки» фильтров

[10] Сверточные слои, в которых прямой и обратный проход алгоритма обратного распространения ошибки переставлены.

RetinaNet [Lin et al., 2017] акцентирует внимание на проблеме несбалансированности обучающей выборки по положительным и отрицательным примерам — выход сетей современных архитектур в большей части описывает отрицательные примеры, поскольку количество предсказываемых описывающих прямоугольников (ОП) велико. Для преодоления этого недостатка предлагается использование функции потерь FL, которая уменьшает вклад «простых» примеров при обучении (к простым примерам относится фон, предсказываемый с высокой вероятностью). FL является модифицированной логистической функцией потерь. При бинарной классификации:

$$FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} \, log(p_t)\,, \tag{30}$$

$$p_t = \begin{cases} p,\ если\ y = 1 \\ 1 - p,\ иначе \end{cases}, \tag{31}$$

где y — эталонная метка, а α_t — «вес» для конкретного примера.

Следует отметить, что существуют также методы, которые используют нейросетевой подход не для обнаружения, а только для извлечения признаков. Так в [Веселов, 2017] для обнаружения объектов на фотоснимках с низким качеством изображения используется признаки, полученные с помощью автоэнкодера, а классификация реализуется с помощью алгоритма RVM [Tipping, 2001].

Следует отметить, что на настоящий момент результаты обнаружения (в отличие от распознавания) даже самыми мощными методами, основанными на СНГО, все еще значительно отстают от возможностей человека даже на специализированных базах данных, подобных COCO и Pascal.

Нейронные сети для сегментации изображения

Сегментация изображений является другой важной задачей компьютерного зрения. До недавнего времени были популярны подходы подобные [Shotton et al., 2008], основанные на использовании случайных лесов. Достаточно неплохие результаты показывали небольшие нейронные сети, применяемые к отдельным небольшим фрагментам изображения [Ciresan et al., 2012]. Однако такой подход

не предоставляет возможности нейронной сети учитывать информацию о контексте тестируемого фрагмента, что очевидным образом накладывает непреодолимые ограничения на дальнейшее повышение качества сегментации без использования сложной последующей обработки.

В настоящее время практически во всех соревнованиях по сегментации изображения убедительно побеждают сверточные нейронные сети, принимающие на вход большой фрагмент изображения или все изображение целиком. В 2014 году был предложен подход [Long et al., 2014] основанный на наблюдении, что сверточные нейронные сети, обученные распознаванию, на картах признаков глубоких сверточных слоев сохраняют пространственную информацию, а карта активаций нейронов имеет тенденцию отражать форму объектов на изображении. Обозначенный эффект демонстрируется на рис. 18.

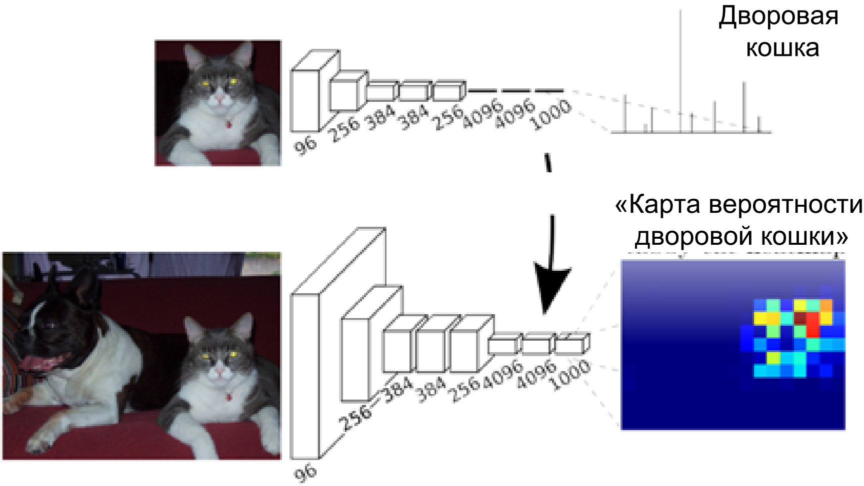


Рис. 18. Иллюстрация свойства сверточных нейронных сетей к пространственному отображению информации об объектах [Long et al., 2014]. Сверху — применение сверточной нейронной сети для распознавания фрагмента изображения, снизу — карта признаков одного из сверточных слоев при предъявлении всего изображения. Высокие активности нейронов топологически соответствуют искомому объекту

При этом в рамках подхода предлагалось использовать полностью сверточную архитектуру без полносвязных слоев. Это позволяет значительно сократить количество весов. Кроме того, как и в задаче обнаружения использовались опорная нейронная сеть для генерации признаков и надстройка, конвертирующая признаки в сегментированные изображения. Во многом наблюдения, сделанные для решения задачи обнаружения, сохраняются и при сегментации. Поскольку на настоящей момент во всех современных архитектурах НС для распознавания используется постоянное понижение пространственного разрешения признакового поля, то применяются так называемые обратные (деконволюционные) сверточные слои, которые представляют собой сверточные слои, в которых прямой и обратный проход алгоритма обратного распространения ошибки переставлены.

Другая известная архитектура сегментации SegNet также использует схожую с [Long et al., 2014] архитектуру и строится в парадигме сверточного автоэнкодера.

Одной из проблем обозначенных подходов является большое количество весов, которым необходимо обучить деконволюционные слои для возможности восстановить карту сегментации до размеров сравнимых с исходным изображением. Уже в статье [Long et al., 2014] было замечено, что возможно использование так называемых редких, расширяющих или прореженных сверток (dilated convolutions), которые помогли в дальнейшем улучшить результаты. Свертка 3×3 с расширяющим шагом p в этом случае делается по формуле:

$$y_{r,c,k} = \sum_{i=-H/2}^{H/2} \sum_{j=-W/2}^{W/2} \sum_{d=1}^{D} x_{p \cdot i+r, p \cdot j+c, d} \cdot w_{i,j,d,k} + b_k, \qquad (32)$$

где (r, c) — это положение фильтра на изображении (или на «карте признаков» предыдущего слоя), $y_{r,c,k}$ — вход нейрона (соответствующий определенному положению фильтра), x — выход нейронов предыдущего слоя, H, W, D — это высота, ширина и глубина фильтра (глубина фильтра, как правило, равна количеству карт признаков на текущем слое или количеству каналов изображения),

$w_{i,j,d,k}$ — это веса фильтра k, полученные в результате обучения, b_k — вес связи с постоянным значением (связь с единицей) (также получается в результате обучения).

Такие свертки позволяют экспоненциально увеличить рецептивные поля нейронов и широко используются в алгоритмах сегментации. В [Chen et al., 2014] использовались условные случайные поля как дополнительный способ улучшить построенную семантическую карту изображения, и это позволило добиться очень серьезного улучшения точности сегментации по сравнению с SegNet. В дальнейшем архитектура была улучшена путем добавления в нее расширяющих сверток и иерархического подхода — формирования карты признаков с помощью фильтров разного размера [Chen et al., 2016].

Существует подход, который рассматривает задачу восстановления сегментированного изображения из карты признаков сверточной нейронной сети, как итеративную задачу последовательного увеличения детализации (очищения — refinement) результатов сегментации. Согласно этому подходу используется специальная архитектура сети RefineNet [Lin et al., 2016], которая способна принимать на вход карту признаков любого размера, а также результат своей работы, но на предыдущем шаге. Очищение начинается с использованием самой высокоуровневой информации– карты признаков с максимальным пространственным сжатием, выход модуля RefineNet в этом случае соединен с входом другого аналогично модуля, но принимающей на вход также карту признаков предыдущего семантического уровня (с большим пространственным разрешением).

Существуют также эффективный подход к сегментации, основанный на использовании архитектуры PSPNet [Zhao et al., 2016], в которую априорно заложена возможность и необходимость осуществлять анализ изображения на разных уровнях разрешения. Это достигается путем применения average pooling с разным размером окна к карте признаков верхнего слоя опорной нейронной сети. После этого полученные уменьшенные карты обрабатываются и вновь приводятся к одному размеру, формируя новую стопку карт признаков, на основании которой строится сегментированное изображение. Архитектура PSPNet приведена на рис. 19.

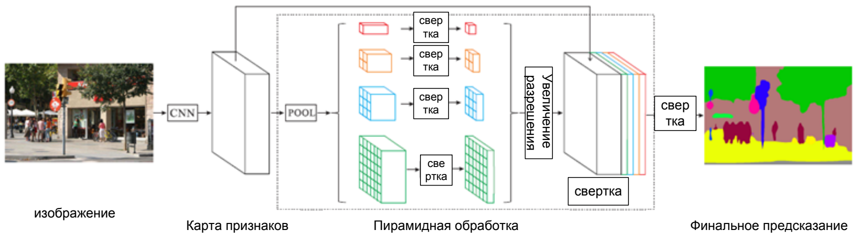


Рис. 19. Архитектура PSPNet

В работе [Луцив, 2015] обращается внимание на то, что обучение нейронной сети сегментации изображения происходит без учета знания об отдельных экземплярах одного класса — предсказывается маска классов объектов и перекрывающиеся объекты одного типа не разделимы тривиальным способом. В связи с этим, было предложено использовать смешанный подход для сегментации и обнаружения. Детекция потенциальных областей интереса для разных объектов происходит с помощью RPN [Ren et al., 2016], после чего в выделенном фрагменте предсказывается маска принадлежности пикселей объекту и маска «непринадлежности» пикселей объекту.

В следующих разделах будут описаны методы обработки изображений, не связанные непосредственно с обнаружением и классификацией объектов, а направленные на улучшение качества изображений, что позволит в дальнейшем более эффективно осуществлять операции детектирования и классификации.

Рассмотренные далее методы улучшения изображений направлены на осуществление следующих операций:

Удаление шумов;

Удаление размытия;

Повышение разрешения;

Колоризация черно-белых изображений;

Перенос стиля для изображений.

Удаление шумов на изображении

Искажение шумами изображений может возникать как на этапе их получения, так и передачи. К возможным причинам следует отнести помехи, присутствующие в канале связи, шум зернистости фото-

материалов и т.д. При эффективном удалении шума предпринимается попытка восстановить исходное изображение, построив его приближение на основе искаженного изображения. При восстановлении изображения требуется минимизировать различия между полученным приближением и оригиналом. Стохастические модели шумовой составляющей во многом определяют сложность решаемой задачи.

Искаженное изображение $g(x, y)$ можно описать как воздействие искажающего оператора H на входное изображение $f(x, y)$ с добавлением шумовой составляющей. При решении задачи восстановления изображения требуется вычислить приближение $\hat{f}(x, y)$ на основе информации об искаженном изображении $g(x, y)$, искажающем операторе H и аддитивном шуме $\eta(x, y)$.

В пространственной области искаженное изображение может быть записано как

$$g(x, y) = h(x, y) * f(x, y) + \eta(x, y), \tag{33}$$

где $h(x, y)$ — функция искажающего оператора в пространственной области, символ «*» используется для обозначения свертки.

К настоящему моменту существует большое число методов подавления шумов на изображении, не связанных с использованием нейронных сетей. Один из подходов основан на использовании сглаживания изображения, например, с помощью медианной фильтрации, когда каждый пиксель изображения заменяется медианным значением из окружающей области пикселей, или с помощью фильтрации Винера [Strela, 2000]. Другой подход связан с разложением изображения на отдельные составляющие на основе вейвлет-преобразования или использования словарей. Вейвлет-преобразование позволяет перевести изображение в частотно-временную область, где его можно более легко отделить от шумовой составляющей, как в методе BLS-SGM [Portilla et al., 2003]. Использование словарей позволяет аппроксимировать зашумленные участки изображения линейной комбинацией из элементов словаря, обученного на изображениях без шума. На основе данного подхода реализован итеративный алгоритм KSVD [Aharon et al., 2006], обучающийся и обновляющий подобный словарь в ходе процедуры подавления шумов.

К последней группе методов следует отнести подходы, использующие глобальную статистику изображений, например, группирующие отдельные участки изображения, схожие визуально, но отличающиеся по пространственному расположению. Подобная группировка используется в методе BM3D [Dabov et al., 2007], который считается одним из наиболее современных и эффективных методов подавления шумов на изображении.

Рассмотренные выше подходы были успешно реализованы на практике, однако все они в том или ином виде используют линейные взаимосвязи для описания своих моделей искажения изображения. Последние исследования показывают, что использование нелинейных моделей, построенных с помощью глубокого обучения, позволит достичь лучших результатов при обработке реальных изображений [Li, 2014].

Один из первых подходов к подавлению шумов на изображениях с применением нейронных сетей глубокого обучения был связан с обучением автоэнкодера [Vincent, 2008]. Под автоэнкодером понимается специальная архитектура искусственных нейронных сетей, позволяющая применять обучение без учителя при использовании метода обратного распространения ошибки.

Возможная схема архитектуры автоэнкодера отражена на рис. 20a.

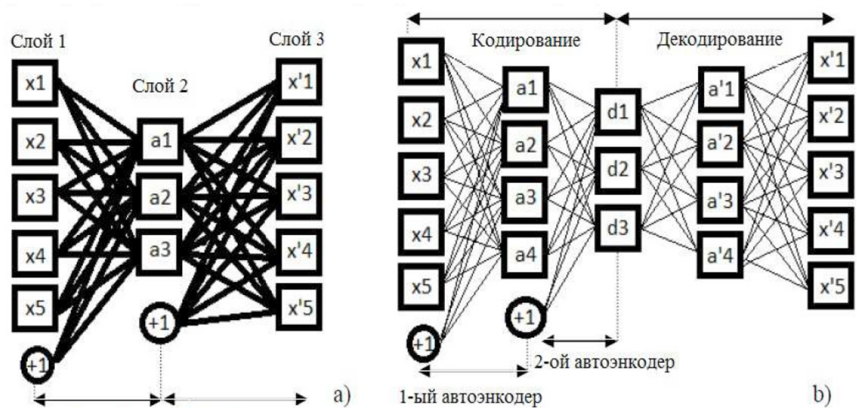


Рис. 20. Схемы архитектур автоэнкодера (*a*) и стекового автоэнкодера (*b*)

Автоэнкодер стремится получить на выходе то же, что и на входе. На рисунке 20a приведен пример архитектуры с полносвязными слоями, однако возможно использование сверток, которые хорошо подходят для обработки изображений. Слои автоэнкодера делятся по типу на кодирующие (понижающие размерность вектора признаков, например, сверточные и подвыборка) и декодирующие (повышающие размерность, например, деконволюционные) [Baldi, 2012].

Автоэнкодер учится выделять характерные черты изображений, чтобы по ним восстановить вход. Также было показано, что использование более сложной стековой архитектуры автоэнкодера (см. рисунок 20b) с большим количеством слоев, позволяет добиться более устойчивой работы алгоритма [Vincent, 2008].

Базовая модель автоэнкодера состоит из функции кодирования $h(\cdot)$, которая сопоставляет входному $x \in R^d$ некоторое скрытое представление $h_x \in R^{d_h}$, и функции декодирования $g(\cdot)$, сопоставляющей скрытое представление обратно реконструированной версии x, таким образом, что $g(h(x)) \approx x$. Параметры автоэнкодера в ходе обучения настраиваются с целью минимизации ошибки построения модели, задаваемой некоторой функцией потерь $L(x, g(h(x)))$. В случае работы с бинарными данными функцией потерь может выступать перекрестная энтропия, для работы с действительными значениями x распространено использование квадратичной ошибки:

$$L(x, g(h(x))) = L_2(x, g(h(x))) = \|x - g(h(x))\|^2. \tag{34}$$

Модель автоэнкодеров, использующаяся для подавления шумов, (Denoising Autoencoders, DAE), вносит небольшие изменения в базовую модель автоэнкодеров. Входные данные искажаются (складываются с шумовой составляющей) перед сопоставлением их со скрытым представлением. DAE обучаются восстановлению исходных данных из их искаженной версии $\tilde{x}$ путем минимизации

$$L(x, g(h(\tilde{x}))) = L_2(x, g(h(\tilde{x}))) = \|x - g(h(\tilde{x}))\|^2. \tag{35}$$

В качестве шумовой составляющей может быть выбран как аддитивный белый гауссовский шум, так и шум, использующий дво-

ичное маскирование. Архитектура энкодера DAE представлена на рис. 21.

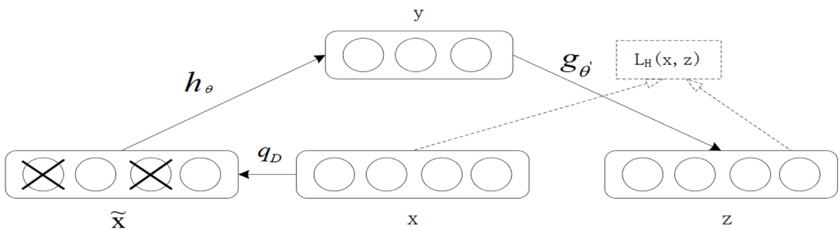


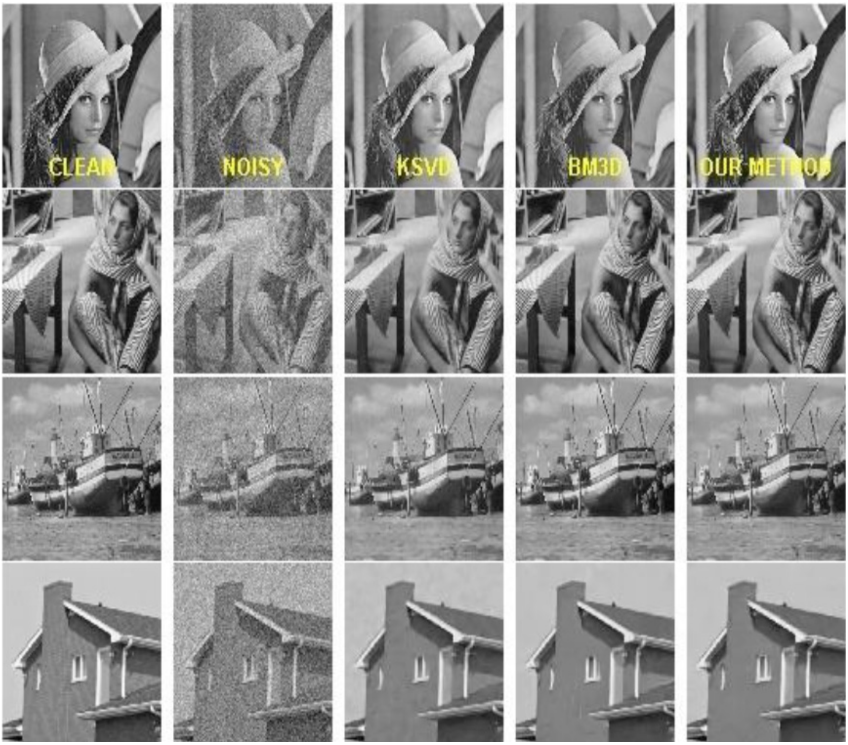
Рис. 21. Архитектура автоэнкодера DAE. Входные данные x искажены с помощью q_D и представляют из себя $\tilde{x}$.

Автоэнкодер сопоставляет искаженные данные $\tilde{x}$ со скрытым внутренним представлением y (с помощью функции кодирования h_θ) и предпринимает попытку восстановить исходное x с использованием функции декодирования g_θ получая реконструкцию z. Ошибка реконструкции оценивается с помощью функции потерь $L_H(x,z)$.

При глубоком обучении DAE автоэнкодеров используется техника послойного обучения [Bengio et al., 2007], когда каждый слой нейронной сети предобучается как DAE. В ходе фазы предобучения при обучении i-го слоя, каждый входной пример сопоставляется со скрытым представлением $h_{i-1}(x)$ и используется как обучающий пример для DAE. Слой обучается заданное число итераций, после чего полученное представление используется в качестве входных данных для следующего слоя. При использовании стековой структуры автоэнкодеров на этапе предобучения веса каждого скрытого слоя настраиваются отдельно. В ходе обучения для более тонкой настройки сети с уже определенными начальными весами используется процедура обратного распространения ошибки.

Результаты подавления шумов с использованием подхода, основанного на стековых автоэнкодерах с глубоким обучением, представлены на рис. 22. В ходе сравнения методов, рассмотренных

в исследовании [Li, 2014], в качестве количественной метрики использовалось пиковое отношение сигнала к шуму на обработанных изображениях. Полученные результаты свидетельствуют о том, что использование нейронных сетей глубокого обучения позволяет реализовать подавление шумов с эффективностью на уровне лучших алгоритмов, применяющих линейные структуры для описания искажающей модели. В среднем метод BM3D, описанный ранее, лучше обрабатывает изображения с регулярной структурой, в то время как нейронные сети показывают лучший результат для сложных структур и более высоких уровней шума.



а *б* *в* *г* *д*

Рис. 22. Сравнение результатов работы методов подавления шума: *а* — исходные изображения, *б* — зашумленные изображения, *в* — метод KSVD, *г* — метод BM3D, *д* — метод, основанный на использовании стековых автоэнкодеров с глубоким обучением

Подходы, основанные на использовании базовой модели автоэнкодеров с полносвязными слоями [Viren, Seung, 2009], также, как и подходы, реализующие стековую архитектуру [Xie et al., 2012], страдают от ограничения, связанного с тем, что для устойчивой работы методов необходимо априорное знание о характере шума, искажающего изображения. Для преодоления данного ограничения был предложен способ, основанный на использовании взвешенной суммы ответов автоэнкодеров, обученных для разных моделей шумов [Agostinelli et al., 2013]. Для выбора оптимальных весов также обучается отдельная нейронная сеть.

Одними из основных ограничений для использования нейронных сетей глубокого обучения для подавления шумов на изображении остаются требования высокой вычислительной мощности для обучения нейронных сетей, а также необходимость наличия больших объемов данных для обучения. В то же время ряд исследований показывает, что при использовании определенной архитектуры сверточной нейронной сети из автоэнкодеров DAE эффективная работа метода становится возможной и при малом размере обучающей выборки (от 300 примеров) [Gondara, 2016].

Удаление размытия на изображении

Удаление размытия (смаза) на изображении состоит в восстановлении исходной резкости изображения, искаженного из-за тряски камеры, движений объектов на снимке или расфокусировки. Это пример обратной задачи, когда значения параметров модели должны быть получены из наблюдаемых данных. Особую сложность при решении вызывает условие неизвестности параметров размытия.

Можно представить, что размытое изображение y получено путем линейного преобразования исходного изображения x с помощью свертки и добавления некоторого шума n:

$$y = k \cdot x + n \quad (36)$$

Задача повышения резкости состоит в восстановлении x, имея только размытое изображение y, без априорного знания о k.

Существует ряд методов, позволяющих убрать неравномерное размытие, вызванное вращением камеры или смещениями в плоско-

сти снимка [Gupta et al., 2010; Hirsch et al., 2011]. Данные методы эффективны при ограничениях параметров движения. Другие подходы направлены на компенсацию размытия при статичном положении камеры и движении объектов в кадре. Они оценивают параметры смаза путем анализа глобальной статистики по изображению [Levin, 2007] или анализа спектра размытия [Chakrabarti et al., 2010]. Еще одно направление связано с совместной оценкой восстановленного изображения и параметров смаза на основе пространственных признаков [Kim, Lee, 2014]. Тем не менее, все данные методы не способны эффективно устранить значительное неравномерное размытие для сложных сцен.

Использование нейронных сетей для повышения резкости позволило разработать алгоритмы как для определения типа размытия и его параметров [Aizenberg et al., 2006; Khare, Nagwanshi, 2011], так и для моделирования размытых изображений [Cho, Don, 1991]. Однако наибольший интерес представляют подходы, позволяющие полностью решить обратную задачу, используя принципы глубокого обучения [Schuler et al., 2014; Sun et al., 2015].

Большая часть подходов, позволяющих реализовать быстрое устранение размытия, состоит из следующих трех операций [Xu et al., 2013]:

1. Выделение признаков. На этом шаге выполняется вычисление представления изображения, которое будет использоваться для оценки параметров размытия. Данное представление может быть найдено путем вычисления градиента размытого изображения на разных масштабах для построения карты наиболее устойчивых контуров [Xu, Jia, 2010] или на основе предварительной оценки восстановленного изображения, рассчитанной на основе нелинейной фильтрации [Cho, Lee, 2009].
2. Оценка параметров размытия. Выделенные признаки используются для расчета параметров размытия.
3. Оценка восстановленного изображения — аппроксимация резкого изображения с использованием рассчитанных параметров.

В работе [Schuler et al., 2014] для выделения признаков и построения поля размытия при движении используется сверточная нейронная сеть. Обучение выполняется таким образом, чтобы для каждого отдельного участка изображения получить вероятность существо-

вания там размытия определенной интенсивности и направления. Таким образом, для каждой области изображения размером 30×30 пикселей выбирается один из возможных 73 векторов размытия. Архитектура нейронной сети, используемой для оценки параметров размытия, представлена на рис. 23.

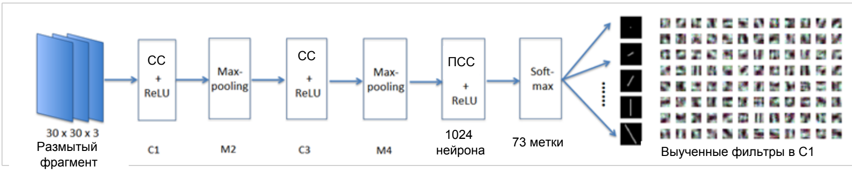
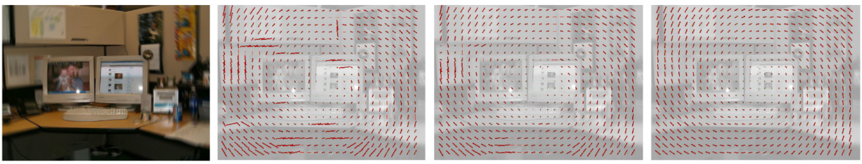


Рис. 23. Архитектура нейронной сети для предсказания параметров размытия (СС – сверточный слой, ПСС — полносвязный слой)

Уточнение результатов предсказания нейронной сети выполняется с помощью марковских случайных полей. После оценок для отдельных участков изображения генерация плотной карты размытия реализована в виде оптимизации модели марковского случайного поля (см. рис. 24). На заключительном этапе работы выполняется восстановление исходного изображения путем решения задачи обратной свертки для неравномерного размытия [Zoran, Weiss, 2011].



а *б* *в* *г*

Рис. 24. Оценка параметров размытия: *а* — размытое изображение, *б* — параметры, найденные нейронной сетью, *в* — параметры, уточненные на основе марковских случайных полей, *г* — эталонные параметры размытия

Описанный подход позволил добиться более точной компенсации, чем существующие методы, не использующие нейронные сети глубокого обучения, однако он применим только для определенных видов движения камеры или объектов в кадре.

Метод, представленный в [Zoran, Weiss, 2011], использует для всех операций единую нейронную сеть и направлен на удаление

размытия любого вида. Сама архитектура нейронной сети состоит из трех модулей. Модуль выделения признаков преобразует изображение в предопределенное градиентное представление, подходящее для оценки ядра фильтра размытия. Модуль оценки параметров размытия объединяет извлеченные признаки и пересчитывает параметры ядра. Модуль оценки восстановленного изображения генерирует текущую версию изображения повышенной резкости. Данные модули объединяются в цепочку и повторяются неоднократно путем увеличения числа слоев в нейронной сети (рис. 25). Это позволяет реализовать данную систему в виде единой нейронной сети, которая реализует итеративное обучение путем обратного распространения ошибки для разных масштабов изображения.

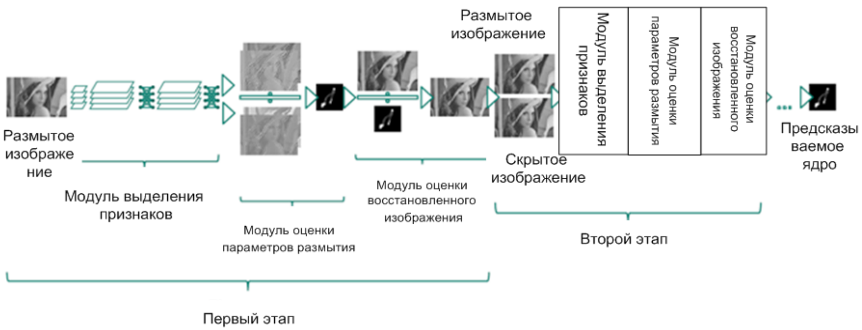


Рис. 25. Структура сверточной нейронной сети, используемой для удаления размытия изображения

Результаты работы описанного метода представлены на рис. 26.



а *б* *в*

Рис. 26. Удаление размытия: *а* — размытое изображение, *б* — метод, не использующий нейронные сети [Cho, Lee, 2009], *в* — метод, использующий сверточные нейронные сети [Schuler et al., 2014]

Количественное и качественное сравнение с существующими аналогами показывает, что метод сопоставим по эффективности с подходами, использующими априорную информацию о виде размытия, и превосходит их, если компенсация производится для определенной, заранее известной, категории изображений.

Хотя данный подход продемонстрировал устойчивую работу, как для разных категорий изображений, так и для размытых изображений с высоким уровнем шума, метод работает нестабильно при больших значениях размытия, что может являться предметом дальнейших исследований.

Повышение разрешения

Разрешение изображения определяется как количество пикселей, составляющих изображение, деленное на единицу площади. Повышение разрешения необходимо для анализа и обработки мелких деталей изображений, что позволяет повысить точность обнаружения и распознавания по сравнению с изображениями с низким разрешением.

Можно выделить два основных подхода к повышению разрешения: с использованием нескольких входных изображений (Multi-Frame Super-Resolution, MFSR) и с использованием только одного исходного изображения (Single Image Super-Resolution, SISR). В данном разделе рассмотрены методы, ориентированные на второй подход.

К настоящему моменту существует большое число методов, направленных на повышение разрешения изображений на основе единственного исходного. К первой группе можно отнести методы, использующие интерполяцию, как способ достижения наилучшей аппроксимации цвета и интенсивности пикселей на основе значений их соседей. Основные алгоритмы интерполяции могут быть отнесены в одну из двух категорий: адаптивные и неадаптивные. Неадаптивные алгоритмы включают в себя метод ближайших соседей, билинейную, бикубическую и сплайновую интерполяции, фильтрацию Ланцоша [Duchon, 1979]. Тогда как неадаптивные алгоритмы обрабатывают все пиксели одинаковым образом, адаптивные методы меняют свои коэффициенты в зависимости от того, что они ин-

терполируют (резкие градиенты и границы, или гладкие текстуры) [Jiang, Moloney, 2002]. Большая часть методов первой группы интерполяторов в той или иной мере страдает от характерных искажений: размытия и граничных ореолов на восстановленном изображении.

Вторая группа методов основана на поиске схожих участков изображения и создании комплексного отображения между изображениями с низким (LR) и высоким (HR) разрешением. В качестве обучающих данных используются пары LR-HR изображений и при восстановлении используются части LR изображений, для которых известны HR аналоги.

В работе [Freeman et al., 2002] применяется предварительно составленный банк участков изображений. На каждой итерации локальные детали объединяются для поиска схожих векторов в базе участков ранее обработанных изображений. Найденные участки высокого разрешения объединяются в результирующее изображение. В исследовании [Glasner et al., 2009] используется избыточность информации, полученной из нескольких схожих частей изображения. Поиск схожих участков производится в пирамиде масштабов (наборе версий исходного изображения разных размеров). Был предложен подход с использованием сверточного разреженного кодирования [Gu et al., 2015], который обрабатывает изображение целиком, а не только отдельные пересекающиеся части, что позволяет улучшить однородность результатов.

К третьей группе можно отнести регрессионные методы, для которых характерен более общий подход к отображению пар участков изображений на основе ядровой регрессии с L2 регуляризацией [Kim, Kwon, 2010]. Это позволяет решить проблему переобучения, возникающую для подходов, повышающих разрешение за счет поиска похожих участков в маломерном множестве обучающей выборки низкого разрешения и совмещения соответствующих участков высокого разрешения для восстановления изображения. Регрессионная формулировка данной задачи может быть решена с помощью гауссовских процессов, решающих деревьев и случайного леса [Schulter et al., 2015].

В настоящее время алгоритмы, основанные на сверточных нейронных сетях, позволили достичь результатов, превосходящих описанные выше методы. В одной из первых работ [Wang et al., 2015]

было реализовано кодирование исходного представления пространства изображений в полносвязную архитектуру нейронной сети, основанной на пороговой обработке и итеративном алгоритме сжатия. Заметное улучшение качества было получено при использовании бикубической интерполяции для увеличения входного изображения и обучения сверточной нейронной сети с тремя слоями [Dong et al., 2014; Dong et al., 2016] (см. рис. 27).

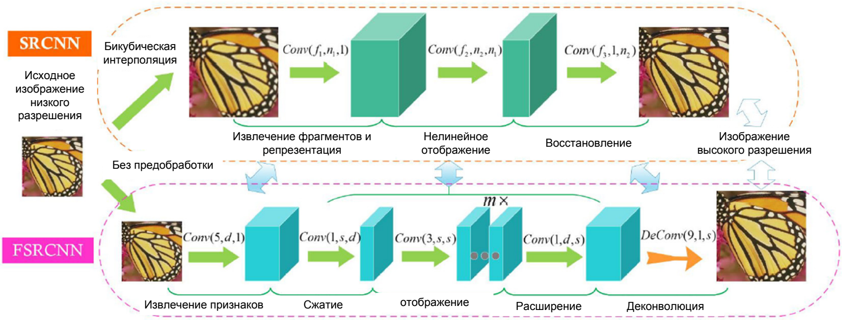


Рис. 27. Архитектура SRCNN (Super-Resolution Convolutional Neural Network) [Dong et al., 2014], обрабатывающая изображение, увеличенное с использованием бикубической интерполяции, и улучшенная архитектура FSRCNN (Fast Super-Resolution Convolutional Neural Network) [Dong et al., 2016], в которой для повышения разрешения задействованы только сверточные слои Было показано, что использование сверточных слоев при повышении разрешения изображения позволяет добиться не только высокой точности, но и сократить время выполнения алгоритмов [Shi et al., 2016].

Обозначим исходное изображение низкого разрешения, которое увеличили до требуемого размера методом бикубической интерполяции как **Y**. Целью повышения разрешения является получение изображения F(**Y**), которое будет наиболее близким к эталонному изображению с высоким разрешением **X**. При использовании сверточных нейронных сетей архитектура сети составляется таким образом, чтобы в ходе обучения было определено отображение F из низкого разрешения в высокое разрешение. Данная процедура состоит из трех основных шагов:

1. Выделение участка изображения и построение представления. Производится извлечение перекрывающихся участков изображе-

ния низкого разрешения **Y** и преобразование информации о данном участке в многомерный вектор. Набор подобных векторов образует карту признаков изображения.

2. Нелинейное отображение. Каждому многомерному вектору нелинейным образом сопоставляется многомерный вектор, соответствующий представлению, выученному нейронной сетью. Полученное сопоставление характеризует участок изображения высокого разрешения.

3. Восстановление. Выполняется объединение отдельных участков изображения высокого разрешения в единое целое.

Реализация данных трех шагов обеспечивает формирование нейронной сети, как представлено на рис. 28.

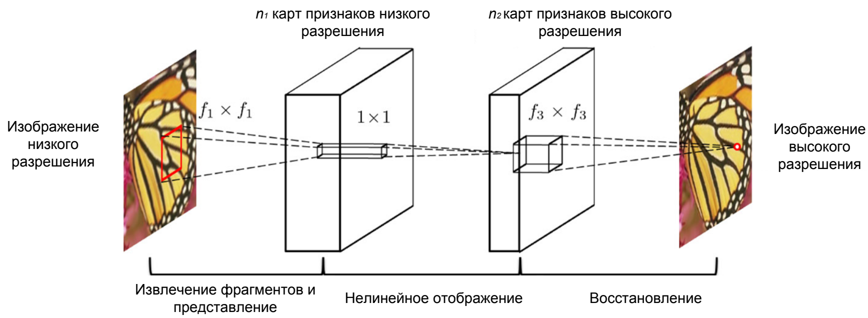


Рис. 28. Общая схема сверточной нейронной сети для повышения разрешения. На вход подается изображение низкого разрешения **Y**, первый сверточный слой извлекает карту признаков изображения. Второй слой нелинейно отображает данные признаки на признаки изображения высокого разрешения. Последний слой объединяет предсказания нейронной сети для отдельных участков изображения для получения результирующего изображения высокого разрешения F(**Y**)

В работе [Kim et al., 2016] представлена сверточная сеть на основе глубокой рекурсии, которая позволяет определить зависимости даже для пикселей изображения, сильно разделенных пространственно, при этом не увеличивая значительно число параметров модели (рис. 29). Следует отметить исследование [Johnson et al., 2016], в котором введена перцепционная функция потерь, позволяющая оптимизировать сеть для восстановления более фотореалистических изображений.

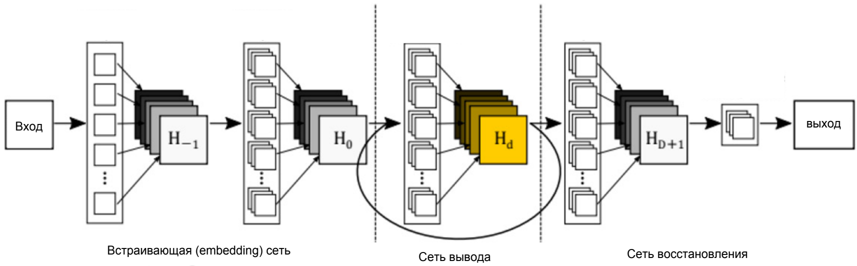


Рис. 29. Схема нейронной сети на основе глубокой рекурсии для повышения разрешения изображений. Часть сети, ответственная за вывод соответствий между участками низкого и высокого разрешения, выделена желтым цветом [Kim et al., 2016].

Подходы, основанные на использовании сверточных нейронных сетей глубокого обучения, наиболее качественно восстанавливают детали изображений, при этом сохраняя высокую скорость работы. Однако использование одной нейронной сети, даже с очень сложной архитектурой, зачастую не позволяет обеспечить фотореалистическое качество с точки зрения человека. Данная проблема может быть решена путем использования конкурирующих нейронных сетей [Ledig et al., 2016].

Конкурирующую нейронную сеть можно задать в виде модели, объединяющей две нейронных сетей: генератор и дискриминатор. Дискриминатор обучается с целью отличать настоящие изображения от результатов работы генератора. В то же самое время в ходе итеративного процесса генератор учится создавать все более реалистичные изображения. В результате обучения сеть-генератор становится способной генерировать изображения, не отличимые от настоящих, согласно критериям, выработанным сетью-дискриминатором.

Модель конкурирующей нейронной сети изначально была предложена для получения реалистичных изображений из матриц случайного шума. При решении задачи повышения разрешения в качестве критерия реалистичности изображения должна использоваться близость по содержанию к изображениям с низким разрешением, поступающим на вход нейронной сети. Изменение

критерия реализуется в виде замены функции потерь для сети-генератора. Предлагаемая перцепционная функция потерь определена как комбинация среднеквадратичной ошибки, рассчитанной на верхних слоях отдельной предобученной нейронной сети, и ошибки дискриминатора [Ledig et al., 2016]. Подобная комбинация позволяет генератору находить решения, неоптимальные согласно критерию среднеквадратичной ошибки, но более реалистические с точки зрения человека.

Результаты работы описанной нейронной сети приведены на рисунке 30. Применение модели конкурирующей нейронной сети позволяет повысить фотореалистичность деталей изображения за счет совместной работы комбинации нейронных сетей и модифицированной функции потерь.



а *б* *в* *г*

Рис. 30. Результаты повышения разрешения: *а* — изображение низкого разрешения, увеличенное до нужного размера бикубической интерполяцией, *б* — результат нейронной сети глубокого обучения, оптимизированной для среднеквадратичной ошибки, *в* — результат конкурирующей нейронной сети, *г* — эталонное изображение высокого разрешения. Разрешение изображения было увеличено в 4 раза.

Колоризация черно-белых изображений

Под колоризацией или раскрашиванием понимается процесс добавления цвета к монохроматическому (черно-белому) изображению или видео. Задача восстановления цвета изображения является недоопределенной, так как одному и тому же монохроматическому изображению могут соответствовать различные цветные изображения.

Существующие подходы к колоризации изображений можно разделить по тому принципу, как используются данные для моделирования соответствия между монохроматическим и цветным изображением. Непараметрические методы при решении данной

задачи используют одно или несколько цветных опорных изображений (которые выбираются из базы изображений автоматически или указываются пользователем), которые служат в качестве источника информации о цвете. Затем, используя принцип поиска аналогов изображения [Hertzmann et al., 2001], цвет переносится из схожих областей опорного изображения на области исходного монохроматического изображения [Welsh et al., 2002; Gupta et al., 2012; Liu et al., 2008]. Для выделения схожих областей при этом обычно используются методы сегментации, основанные на выделении границ или использующие представление изображения в виде взвешенного графа.

Параметрические методы основаны на обучении функции, предсказывающей цвет, с использованием большого объема цветных изображений. В ходе обучения данная проблема формулируется либо в виде регрессионной задачи в непрерывном цветовом пространстве [Deshpande et al., 2015], либо как задача классификации дискретных цветовых значений [Charpiat et al., 2008].

Использование сверточных нейронных сетей можно отнести к группе параметрических методов, так как нейронная сеть также обучается классифицировать отдельные цвета, однако больше размеры как обучающей выборки, так и самой модели отображения монохроматического изображения в цветное. Существует ряд реализаций, основанных на сверточных нейронных сетях [Larsson et al., 2016; Iizuka et al., 2016; Hariharan et al., 2015], которые отличаются видом функции потерь и архитектурой самой сети. В данном разделе будет рассмотрена нейронная сеть, показавшая наилучшие результаты колоризации [Zhang et al., 2016], ее архитектура представлена на рис. 31.

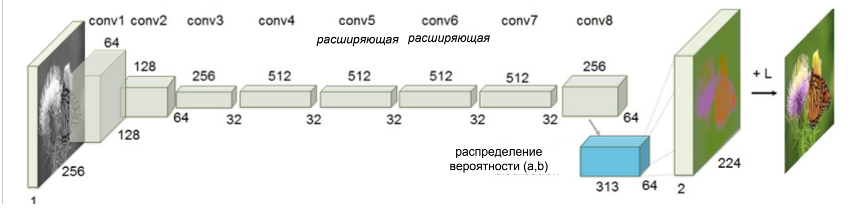


Рис. 31. Архитектура сверточной нейронной сети, решающей задачу колоризации

В рамках данного подхода используется цветовое пространство CIE Lab. Задача формулируется следующим образом: на вход подается канал светлоты (lightness, L) $\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$, необходимо обучиться отображению $\hat{\mathbf{Y}} \in F(\mathbf{X})$ для двух каналов с информацией о цвете $\mathbf{Y} \in \mathbb{R}^{H \times W \times 2}$, где H, W — размеры изображения.

Целевой функцией, подлежащей оптимизации, является евклидова норма разности между предсказанным $\hat{\mathbf{Y}}$ и истинным $\mathbf{Y}$ значением цвета:

$$\mathrm{L}_2(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2} \sum_{h,w} \left\| \mathbf{Y}_{h,w} - \hat{\mathbf{Y}}_{h,w} \right\|_2^2. \quad (37)$$

Однако данная целевая функция неустойчива из-за внутренней неоднозначности и мультимодальной сути задачи колоризации. Использование евклидовой метрики приводит к усреднению цветовых значений и блеклым, обесцвеченным результатам. Для преодоления данного недостатка задача формулируется в виде мультиноминальной (многоклассовой) классификации. Производится квантование выходного цветового пространства *ab* на отдельные ячейки, как показано на рис. 32.

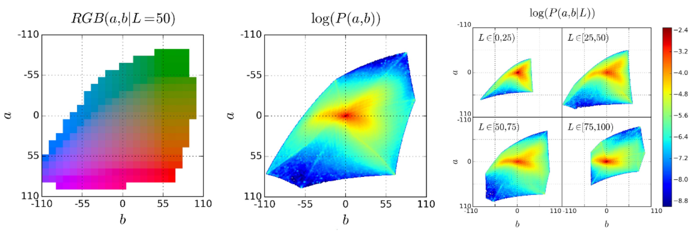


Рис. 32. Квантованное цветовое пространство *ab* (шаг сетки 10)

Для заданного $\mathbf{X}$ ищется отображение $\hat{\mathbf{Z}} = G(\mathbf{X})$ на распределение вероятности по возможным цветам $\hat{\mathbf{Z}} \in [0,1]^{H \times W \times Q}$, где Q — число квантованных значений *ab*. На последнем этапе работы нейронной сети данное распределение вероятности отображается на действительные цветовые значения $\hat{\mathbf{Y}}$. Результаты колоризации черно-белых фотографий с использованием описанного алгоритма представлены на рис. 33.



Рис. 33. Результаты колоризации черно-белых фотографий с применением сверточных нейронных сетей

Полученные результаты указывают на то, что использование нейронных сетей глубокого обучения для колоризации позволяет эффективно решать данную задачу, однако могут возникать сложности при восстановлении цвета для изображений, характер которых резко отличается от тех, что входили в состав обучающей выборки.

Перенос стиля для изображений

Под переносом стиля понимается техника перерисовки изображения с использованием цветов и текстур одного изображения (задающего стиль), но с сохранением содержания исходного изображения. Данное направление исследований появилось именно в связи с развитием сверточных нейронных сетей глубокого обучения, как отмечено в пионерской работе [Gatys et al., 2015].

Сверточная нейронная сеть для решения задачи распознавания объектов учится выделять на основе исходных изображений те признаки, которые определяют содержание изображения и все меньше зависят от конкретных значений интенсивности, подверженных искажению из-за изменений освещенности, наличию шумов и т. п. Используя слои обученной нейронной сети, близкие к выходному слою, становится возможным разделить высокоуровневое содержание изображения (объекты и их взаимное расположение в кадре) и стилистические особенности (цветовая гамма и текстура). Пример реконструкции двух составляющих изображения, содержательной и стилистической, с помощью сверточной нейронной сети представлен на рис. 34.

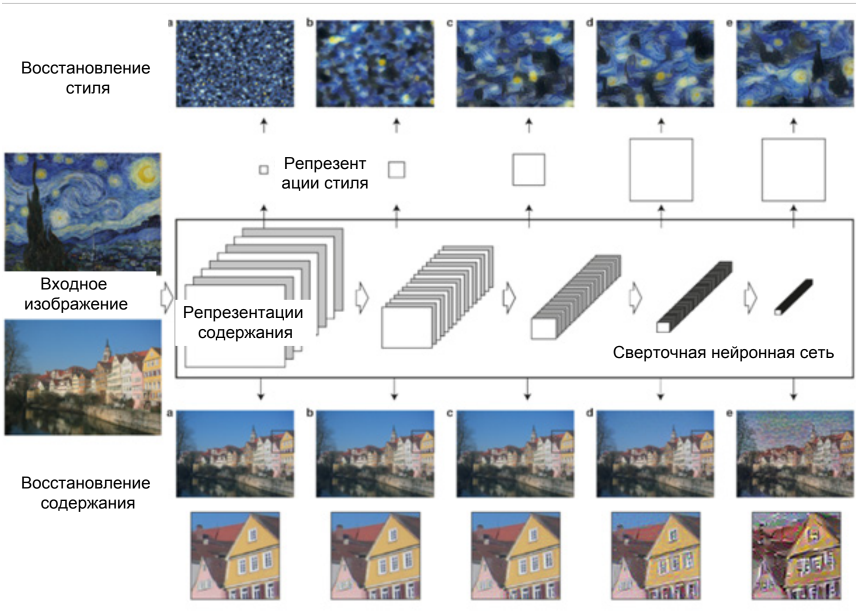


Рис. 34. Использование сверточной нейронной сети для восстановления содержательной и стилистической составляющей. Два входных изображения слева представлены в виде набора изображений, полученных на основе свертки с фильтрами определенных слоев обученной нейронной сети (слои расположены слева направо от первых к последним и обозначены буквами *a*–*e*) [Gatys et al., 2015]. В нижней части представлен процесс получения содержания изображения, нижние слои a–c позволяют получить восстановленное изображение без заметных искажений. В верхней части представлен процесс извлечения стиля входного изображения, верхние слои *d*–*e* позволяют получить информацию о цвете и текстуре, отбросив содержание

Задача переноса стиля изображения может быть сформулирована как оптимизационная задача. Требуется получить изображение **X**, которое будет отличаться как можно меньше от изображения **c** по содержанию и в то же время отличаться как можно меньше по стилю от изображения s:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}}(\alpha L_{content}(\mathbf{c},\mathbf{X}) + \beta L_{style}(\mathbf{s},\mathbf{X})), \quad (38)$$

где $L_{content}$ — функция потерь для содержания изображения, L_{style} — функция потерь для стиля изображения, α, β — весовые коэффициенты.

В качестве функции потерь для содержания изображения используется взвешенная сумма невязок между входным изображением и результирующим, в то время как функция потерь для стиля изображения определяется путем аппроксимации распределения признаков на заданных слоях нейронной сети многомерным нормальным распределением с применением матрицы Грама [Johnson et al., 2016].

Результаты переноса стиля, полученные при использовании сверточных нейронных сетей, представлены на рис. 35.



Рис. 35. Пример переноса стиля картины на изображение естественной сцены с применением нейронных сетей глубокого обучения

Одним из главных недостатков описанного алгоритма является высокая требовательность к вычислительным ресурсам и низкое быстродействие. Для преодоления данного недостатка был предложен подход, основанный на оптимизации не результирующего изображения, а нейронной сети, уже настроенной для передачи определенного стиля [Johnson et al., 2016]. Оптимизационный процесс выполняется лишь один раз для обучения сети-генератора стиля, что позволило повысить быстродействие на несколько порядков (генерация одного изображения теперь измеряется десятками миллисекунд).

Заключение

В данной главе приведен обзор и анализ современных решений в области автоматического анализа изображений с помощью нейронных сетей глубокого обучения, которые на настоящий момент являются наиболее успешными алгоритмами для распознавания, сегментации, обнаружения и улучшения изображений.

Среди рассмотренных задач наиболее успешно решена задача распознавания. В связи с этим, многие модели нейронных сетей-классификаторов выступают в качестве базовых «экстракторов признаков» для решения других задач. Не так давно появились решения для обнаружения объектов, которые не требуют многократного запуска сети-классификатора на разных фрагментах кадра, а позволяют осуществлять регрессию изображения в набор описывающих прямоугольников за один проход. Для задачи сегментации изображений на настоящий момент используют полностью сверточные сети (в которых отсутствуют полносвязные слои). Алгоритмы обработки изображений, использующие искусственные нейронные сети, превосходят имеющиеся аналоги по точности и в ряде случаев позволяют добиться фотореалистического качества результирующих изображений. Это можно объяснить тем, что сверточные сети во многом заимствовали особенности обработки изображений в естественных нейронных архитектурах и поэтому обеспечивают результаты, согласующиеся с особенностями зрительного восприятия человека.

Помимо высокого качества при решении рассмотренных задач анализа изображений к несомненным плюсам можно отнести про-

стоту модели идеального нейрона и возможность организации массово параллельных вычислений, что позволяет эффективно задействовать архитектуры искусственных нейронных сетей в решениях на основе программируемых логических интегральных схем и графических ускорителей.

Литература

Байдаков М. П., Красильников Н. Н., Пастухов О. В. О различении человеком изображений в гауссовом шуме // Автометрия 1973. N1. С. 7–14

Веселов А. Алгоритм детектирования объектов на фотоснимках с низким качеством изображения // Программные продукты и системы, 2017, Т. 1, № 30. С. 130–136

Красильников Н. Н. Теория передачи и восприятия изображений. — М.: Радио и связь, 1986. — 247 с.

Луцив В. Р. Нейронные сети глубокого обучения // Оптический журнал. Т. 82, № 8. 2015. С. 11–23

Малашин Р. О., Кадыков А. Б. Исследование обобщающих способностей сверточных нейронных сетей при формировании признаков, инвариантных к вращению // Оптический журнал — 2015. — Т. 82. — No 8. — С. 24–32.

Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов. пер. с англ. М.: Радиоисвязь, 1987. 400 с.

Agostinelli, F., Anderson M. R., Lee H. Adaptive multi-column deep neural networks with application to robust image denoising // Advances in Neural Information Processing Systems. 2013.

Aharon M., Elad M., Bruckstein A. K-svd: An algorithm for designing over complete dictionaries for sparse representation // IEEE Transactions on Signal Processing (TIP). 2006. Vol. 54. № 11. P. 4311–4322.

Aizenberg I., Paliy D., Moraga C., Astola J. Blur identification using neural network for image restoration // Computational Intelli-

gence, Theory and Applications. Springer. 2006. Vol. 38. P. 441–455.

Baldi P. Autoencoders, Unsupervised Learning, and Deep Architectures. — Department of Computer Science, University of California, Irvine, 2012.

Bengio Y., Lamblin P., Popovici D., et al. Greedy layer-wise training of deep networks // Advances in neural information processing systems. 2007. Vol. 19. P. 153–172.

Caffe Deep Learning Library. URL: http://caffe.berkeleyvision.org/ (accessed 10.04.2017).

Chakrabarti A., Zickler T., Freeman W. Analyzing spatially-varying blur // CVPR. 2010. P. 2512–2519.

Charpiat G., Hofmann M., Scholkopf B. Automatic image colorization via multimodal predictions // Computer Vision ECCV. 2008. P. 126–140.

Chen L., Papandreou G., Kokkinos I., Murphy K., Yuille A. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. 2016. URL: https://arxiv.org/abs/1606.00915

Chen L., Papandreou G., Kokkinos I., Murphy K., Yuille A. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. 2014. URL: https://arxiv.org/abs/1412.7062

Cho C. M., Don H. S. Blur identification and image restoration using a multi-layer neural network // IEEE Int. Joint Conf. Neural Networks. 1991. P. 2558–2563.

Cho S., Lee S. Fast motion deblurring // ACM Trans. Graphics 28.5. 2009 P. 145.

Ciresan D. Giusti A., Gambarderlla L., J. Shmidhuber. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. 2012

Dabov K., Foi A., Katkovnik V., Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering // IEEE Transac-

tions on Image Processing (TIP). 2007. Vol. 16. № 8. P. 2080–2095.

Deshpande A., Rock J., Forsyth D. Learning large-scale automatic image colorization // Proceedings of the IEEE International Conference on Computer Vision. 2015. P. 567–575.

Dong C., Loy C., He K. Learning a deep convolutional network for image super-resolution // European Conference on Computer Vision. 2014. P. 184–199.

Dong C., Loy C. C., He K., Tang X. Image super-resolution using deep convolutional networks // IEEE transactions on pattern analysis and machine intelligence. 2016. Vol. 38. P. 295–307.

Duchon C. E. Lanczos filtering in one and two dimensions // Journal of Applied Meteorology. 1979. Vol. 18. P. 1016–1022.

Freeman W., Jones T., Pasztor E. Example-based super-resolution // IEEE Computer graphics and Applications. 2002. Vol. 22. № 2. P. 56–65.

Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position // Biological Cybernetics. 1980. V. 36. № 4. P. 93–202.

Gatys L., Ecker A., Bethge M. A neural algorithm of artistic style // Nature Communications. 2015.

Girshick R. Fast R-CNN. 2015. URL: arXiv:1504.08083.

Girshick R., Donahue J. , Darrell T. , and Malik J. Region-based convolutional networks for accurate object detection and segmentation. TPAMI, V.38, № 1. 2015. P 142–158.

Glasner D., Bagon S., Irani M. Super-resolution from a single image // Computer Vision. 2009. P. 349–356.

Gondara, L. Medical image denoising using convolutional denoising autoencoders // 2016. arXiv preprint arXiv:1608.04667.

Gu S., Zuo W., Xie Q., Meng D., Feng X., Zhang L. Convolutional sparse coding for image super-resolution // Proceedings of the IEEE International Conference on Computer Vision. 2015. P. 1823–1831.

Gupta A., Joshi N., Zitnick C., Cohen M., Curless B. Single image deblurring using motion density functions // ECCV. 2010.

Gupta R. K., Chia A. Y.S., Rajan D., Ng E. S., Zhiyong H. Image colorization using similar images // Proceedings of the 20th ACM international conference on Multimedia. 2012. P. 369–378.

Hariharan B., Arbelaez P., Girshick R., Malik J. Hypercolumns for object segmentation and fine-grained localization // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 447–457.

He K., Zhang X, Ren S., Sun J. Deep residual learning for image recognition. 2015. URL: arXiv:1512.03385 (accessed 2017.07.03).

He K., Zhang X., Ren S., Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification // 2015. URL: https://arxiv.org/pdf/1502.01852.pdf.

Hertzmann A., Jacobs C. E., Oliver N., Curless B., Salesin D. H. Image analogies // Proceedings of the 28th annual conference on Computer graphics and interactive techniques. 2001. P. 327–340.

Hirsch M., Schuler C., Harmeling S., Scholkopf B. Fast removal of non-uniform camera shake // ICCV. 2011.

Hosseini H., Xiao B. and Poovendran R. Google's Cloud Vision API Is Not Robust to Noise. 2017. URL: https://arxiv.org/pdf/1704.05051.pdf

Huang G., Liu Z., Weinberger K. Q., and van der Maaten L. Densely connected convolutional networks // CVPR. 2017.

Iizuka S., Simo-Serra E., Ishikawa H. Let there be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification // ACM Transactions on Graphics (Proc. of SIGGRAPH 2016). 2016. Vol. 35 (4).

ImageNet Large Scale Visual Recognition Challenge 2012 [official website]. URL: http://www.image-net.org/challenges/LSVRC/2012/ (accessed: 04.01.2014).

Ioffe S. and Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift // ICML. 2015.

Jiang H., Moloney C. A new direction adaptive scheme for image interpolation // International Conference on Image Processing 3. 2002. P. 369–372.

Johnson J., Alahi A., Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution // European Conference on Computer Vision. 2016. P. 694–711.

Johnson J., Alahi A., Li F. Perceptual losses for real-time style transfer and super- resolution // European Conference on Computer Vision (ECCV). 2016. P 694–711.

Karen Simonyan, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. 2014 // URL: arXiv:1409.1556 (accessed 2017.07.03).

Khare C., Nagwanshi K.K. Image Restoration in Neural Network Domain using Back Propagation Network Approach // Int. J. Computer Information Systems 2.5. 2011. P. 25–31.

Kim J., Kwon Lee J., Mu Lee K. Deeply-recursive convolutional network for image super-resolution // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. P. 1637–1645.

Kim K.I., Kwon Y. Single-image super-resolution using sparse regression and natural image prior // IEEE transactions on pattern analysis and machine intelligence. 2010. Vol. 32. P. 1127–1133.

Kim T.H., Lee K.M. Segmentation-free dynamic scene deblurring // CVPR. 2014.

Krizhevsky A., Sutskever I., Hinton G. ImageNet Classification with Deep Convolutional Neural Networks // Advances in Neural Information Processing. 2012. V. 25. P. 1097–1105.

Larsson G., Maire M., Shakhnarovich G. Learning representations for automatic colorization // European Conference on Computer Vision. 2016.

Laws K. Rapid texture identification // SPIE Image Processing for Missile Guidance. 1980. V. 238. P. 376–380.

LeCun Y. et al. Backpropagation Applied to Handwritten Zip Code Recognition // Neural Computation. 1989. Vol. 1. Issue 4. P. 541–551.

LeCun Y., Bottou L., Bengio Y., and Haffner P. Gradient-based learning applied to document recognition // *Proceedings of the IEEE*, november 1998.

Ledig C., Theis L., Husz, Caballero J., Cunningham A., Acosta A., Aitken A., Tejani A., Totz J., Wang Z. Photo-realistic single image super-resolution using a generative adversarial network // 2016. arXiv preprint arXiv:1609.04802.

Levin A. Blind motion deblurring using image statistics // NIPS. 2007.

Li H. Deep Learning for Image Denoising // International Journal of Signal Processing, Image Processing and Pattern Recognition. 2014. Vol. 7. № 3. P. 171–180.

Lin T., Goyal P., Girshick R., He K., Dollár P. Focal Loss for Dense Object Detection. 2017. URL: https://arxiv.org/abs/1708.02002

Lin, G., Milan, A., Shen, C., Reid, I. RefineNet: Multi-Path Refinement Networks for High-Resolution Semantic Segmentation. 2016. URL: https://arxiv.org/abs/1611.06612

Liu W., D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.—Y. Fu, A.C. Berg. SSD: Single Shot MultiBox Detector // ECCV. 2015

Liu X., Wan L., Qu Y., Wong T.T., Lin S., Leung C.S., Heng P. A. Intrinsic colorization // ACM Transactions on Graphics (TOG). 2008. Vol. 27. P. 152.

Long J., Shelhamer E., Darrell T. Fully Convolutional Networks for Semantic Segmentation. 2014. URL: https://arxiv.org/abs/1606.00915.

Malashin R. Matching of Aerospace Photographs with the use of Local Features // Journal of Physics: Conference Series — 2014, Vol. 536, No. 1, pp. 012018.

Malashin R. O., Anisimov N. A. Training Deep Neural Network for Accurate Age and Gender Recognition on Small Set of Data // Proceedings of SPCN2016–2016, pp. 19–22.

Mikolov T., Karafiat M., Burget L., Cernocky J., Khudanpur S. Recurrent neural network based language model // 11th Annual Conference of the International Speech Communication Association. Japan. 2010. P. 1045–1048.

Min Lin, Qiang Chen, Shuicheng Yan. Network In Network // 2013. URL: https://arxiv.org/pdf/1312.4400.pdf (accessed 2017.07.03).

Portilla J., Strela V., Wainwright M. J., Simoncelli E. P. Image denoising using scale mixtures of Gaussians in the wavelet domain // Image Processing. 2003. Vol. 12. № 11. P. 1338–1351.

Redmon J., Divvala S. Giirshick R., Farhadi A. You only look once: unified, real-time object detection. 2015. URL: arXiv:1506.02640. 2016.

Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger. 2016. arXiv:1612.08242v1.

Rumelhart E., Hinton G., Williams R. Learning internal representations by error propagation. ICSReport. 1985.

Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. 2016, URL: arxiv.org/abs/1506.01497.

Schmidhuber J. Deep Learning in Neural Networks: An Overview // Neural Networks. 2015. Vol. 61. P. 85–117.

Schuler C. J., Hirsch M., Harmeling S., Scholkopf B. Learning to deblur // 2014. arXiv preprint arXiv:1406.7444.

Schulter S., Leistner C., Bischof H. Fast and accurate image upscaling

with super-resolution forests // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 3791–3799.

Shen, Zhiqiang and Liu, Zhuang and Li, Jianguo and Jiang, Yu-Gang and Chen, Yurong and Xue, Xiangyang. DSOD: Learning Deeply Supervised Object Detectors from Scratch // ICCV.2017

Shi W., Caballero J., Husz, Totz J., Aitken A.P., Bishop R., Rueckert D., Wang Z. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. P. 1874–1883.

Shotton J., Johnson M., Cipolla R. Semantic Texton Forests for Image Categorization and Segmentation// In Computer Vision and Pattern Recognition, 2008. CVPR2008. IEEE Conference on, pages 1–8. IEEE, 2008.

Strela V. Denoising via block Wiener filtering in wavelet domain // 3rd European Congress of Mathematics. Barcelona. 2000.

Sun J., Cao W., Xu Z., Ponce J. Learning a convolutional neural network for non-uniform motion blur removal // IEEE CVPR. 2015. P. 769–777.

Szegedy C. et al. Going deeper with convolutions. 2014, URL: arXiv:1409.4842 (accessed 2017.05.03).

The PASCAL Visual Object Classes [official websize]. URL: http://pascallin.ecs.soton.ac.uk/challenges/VOC/ (accessed: 17.03.2014).

Tipping, Michael E. Sparse Bayesian Learning and the Relevance Vector Machine // Journal of Machine Learning Research, 2001, № 1. P. 211–244.

Vincent P., Larochelle H., Bengio Y., Manzagol P. Extracting and composing robust features with denoising autoencoders // In Proc. Int. Conf. Mach. Learn. 2008. P. 10961103.

Viren J., Seung S. Natural image denoising with convolutional networks // Advances in Neural Information Processing Systems. 2009.

Wang Z., Liu D., Yang J., Han W., Huang T. Deep networks for image super-resolution with sparse prior // Proceedings of the IEEE International Conference on Computer Vision. 2015. P. 370–378.

Welsh T., Ashikhmin M., Mueller K. Transferring color to greyscale images // ACM Transactions on Graphics (TOG). 2002. Vol. 21 (3). P. 277–280.

Xie J., Linli X., Enhong C. Image denoising and inpainting with deep neural networks // Advances in Neural Information Processing Systems. 2012.

Xu L., Jia J. Two-Phase Kernel Estimation for Robust Motion Deblurring // Computer Vision — ECCV 2010. Lecture Notes in Computer Science. Springer. 2010. P. 157–170.

Xu L., Zheng S., Jia J. Unnatural l0 sparse representation for natural image deblurring // IEEE Conf. Computer Vision and Pattern Recognition. 2013. P. 1107–1114.

Zeiler M.D., Fergus R. Visualizing and understanding convolutional networks // European Conference on Computer Vision (ECCV). Part I. September 6–12, 2014. Zurich, Switzerland. P. 818–833.

Zeiler M.D., Krishnan D., Taylor G.W., Fergus R. Deconvolutional networks // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 13–18 June 2010. San Francisco, USA. P. 2528–2535.

Zhang R., Isola P., Efros A.A. Colorful image colorization // ECCV. 2016.

Zhao H., Shi J., Qi X., Wang X., Jia J. Pyramid Scene Parsing Network. 2016. URL: https://arxiv.org/abs/1612.01105.

Zoran D., Weiss Y. From learning models of natural image patches to whole image restoration // ICCV. 2011.

Глава 13.
Обработка зрительной информации в искусственных и биологических нейронных сетях

Е. Ю. Малахова

В последние годы искусственные нейронные сети стали широко использоваться для решения задач, связанных с обработкой изображений: распознавание объектов и текста, идентификация лиц, семантическая разметка, художественная стилизация и прочее.

В отличии от алгоритмов компьютерного зрения предыдущего поколения, нейронные сети эффективно работают с естественными изображениями, подобными тем, что встречаются человеку ежедневно, и не требуют ручного программирования признаков: то, какие характеристики изображения выделять, и как они соотносятся с конечной задачей, определяется автоматически во время обучения модели. Нейронные сети также обладают возможностью работать с образами независимо от того, при каком освещении они были запечатлены, с какого ракурса и сколько пространства на снимке занимают, что долгое время считалось сложнейшей задачей. Многообразие форм и представлений даже самых простых объектов не позволяло составить список правил, по которым бы можно было отличать эти объекты от других, в чем-то схожих с ними, при постоянно меняющихся условиях наблюдения. Внутренняя структура и качество конечного результата работы нейронных сетей позволяют рассматривать их как возможную модель зрительной системы человека, изучать происходящие процессы и проводить двусторонние аналогии между указанными системами.

Репрезентация образов в зрительной системе

Зрительная система человека характеризуется удивительной точностью в распознавании образов за кратчайшее время и несмотря на вариации в их размерах, положении и освещении. Считается [DiCarlo et. al., 2007], что основной задачей зрительной обработки является формирование нейрональных репрезентаций изображений,

устойчивых к таким трансформациям, вызванным постоянно меняющимися условиями окружающей среды. Эта способность называется инвариантным распознаванием образов.

Обеспечение инвариантного распознавания образов в коре головного мозга является ключевым аспектом восприятия и широко исследуется в нейрофизиологии и когнитивных науках. На Рис. 1 изображено иерархическое строение вентрального зрительного пути. Сигнал от первичной зрительной коры, где нейронами выделяются простые признаки в локальных областях (такие как ориентация линий, частота, цвет), распространяется далее с повышением сложности выделяемых признаков, вплоть до нижневисочной коры, отвечающей за комплексные образы (такие как человек, животные, инструменты, и т.д.) [Sato et al., 2013]. Нейроны в нижневисочной коре охватывают значительную часть всего зрительного поля.

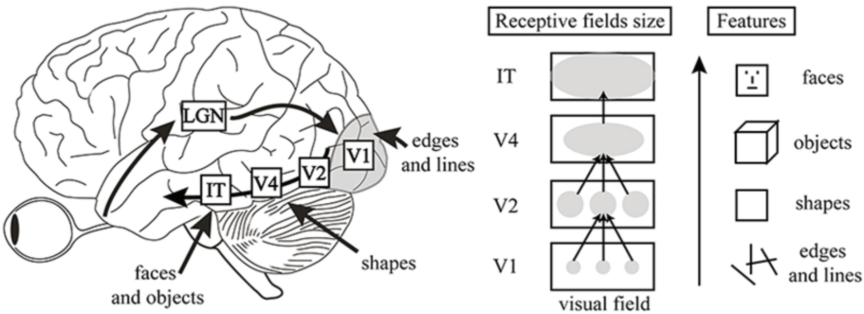


Рис. 1. **Слева:** Иерархическое строение зрительной системы с прямым распространением сигнала: от сетчатки к латеральному коленчатому телу (LGN), далее к областям зрительной коры V1, V2, V4, и нижневисочной (IT) коре. Принятие решения на основании поступившего сигнала осуществляется в лобной коре. **По центру:** Нейроны на нижних уровнях иерархии имеют маленькие рецептивные поля (обрабатывают сигнал с малой части изображения), в то время как нейроны в более высоких областях видят все большую часть зрительного поля, аккумулируя сигналы с нижестоящих уровней. **Справа:** Нижние уровни иерархии отвечают за распознавание базовых признаков изображений, таких как углы и линии. Нейроны высоких уровней собирают с нижележащих областей и кодируют информацию о более сложных признаках, таких как сложные формы, объекты и лица. Изображение взято из Манасси и др. [Manassi et al., 2013].

Так как процесс распознавания образов осуществляется на высоких уровнях иерархии зрительной системы, то его не так просто соотнести с поступающим зрительным сигналом, что затрудняет изучение: поступающая информация обрабатывается с последовательным приращением сложности. Для понимания того каким образом кодируется и хранится информация о зрительном сигнале, какие признаки являются универсальными и необходимыми для успешного решения задачи распознавания объектов, в нейрофизиологии применяются различные методы регистрации нейрональной активности, такие как электроэнцефалография, функциональная магнитно-резонансная томография, запись ответов отдельных нейронов при помощи вживленных в кору микроэлектродов. По результатам исследований выделяется группа стимулов, реакция на которые в определенной изучаемой области существенно превышает уровень спонтанной активности.

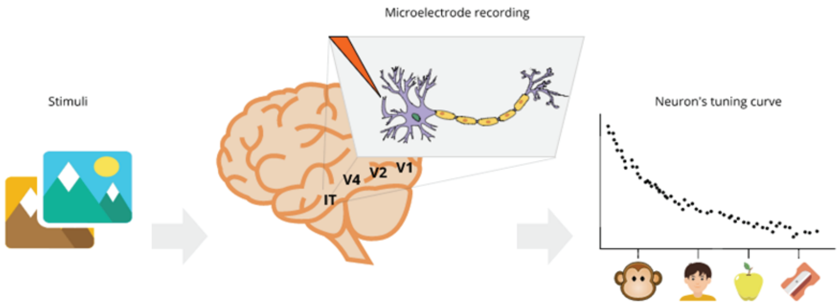


Рис. 2. Изучение функции нейрона на основе данных, полученных с микроэлектрода, записывающего ответ на предъявляемые стимулы. По результатам эксперимента формируется график настройки нейрона — проранжированные по уровню активации стимулы

Стимулы могут быть не только выделены в отдельную группу, но проранжированы от наиболее активирующих к ингибирующим для построения кривой настройки (tuning curve) нейрона (Рис. 2).

Репрезентация образов в искусственных нейронных сетях

Для работы с изображениями используется особенный тип нейронных сетей — сверточные сети, архитектура которых была предло-

жена в 80-х годах прошлого века Кунихико Фукушимой [Fukushima, 1980]. Фукушима интересовался организацией высших функций в мозге человека и на основе открытий Д. Хьюбела и Т. Визела о процессах в зрительной коре разработал модель искусственной нейронной сети «Неокогнитрон» для задачи распознавания рукописных цифр. Эта модель стала прототипом для последующих усовершенствований в архитектуре и методах обучения, которые на сегодняшний день достигают качества распознавания на уровне человеческих возможностей.

Искусственные нейронные сети позволили реализовать инвариантное распознавание в программном виде [LeCun, 2004]. Неоднократно было показано сходство в работе сетей с процессами, происходящими в вентральном зрительном пути. Например, в отличие от всех предыдущих биологически правдоподобных моделей, глубокие нейронные сети помимо качества распознавания объектов также хорошо предсказывают ответ нейронов нижневисочной коры [Cadieu et. al., 2014]. Указанное свойство, хотя и не говорит об идентичности протекающих внутри процессов, но уже свидетельствует о том, что искусственные сети не ориентируются только лишь на метрику качества, а реализуют в чем-то сходные с биологическими системами принципы.

Исследования работы областей вентрального пути показали сходство выделяемых на разных уровнях признаков с теми, что формируются в искусственных нейронных сетях [Güçlü, Gerven, 2015].

Помимо решения практических задач, сверточные нейронные сети рассматриваются исследователями как модель зрительной системы и, в отличие от биологических сетей, гораздо проще поддаются изучению. Были разработаны методы визуализации, позволяющие изобразить «идеальный» стимул для каждого нейрона [Simonyan, Vedaldi, Zisserman, 2013], а также выбрать из предложенных изображений, те, которые вызывают максимальную активацию [Yosinski et. al., 2015].

Возьмем в качестве объекта рассмотрения нейронную сеть, которая была обучена задаче определять есть ли текст на фотографии. Архитектура сети соответствует классической архитектуре AlexNet. Для того, чтобы посмотреть что находится внутри обученной нейро-

сети, мы визуализируем работу некоторых слоев во время прохождения через них изображений.

Приведенные ниже данные можно разделить на две категории: те, которые специфичны для конкретного изображения и те, что являются постоянной составляющей сети, в независимости от того, какое изображение обрабатывается. Так, фильтры слоев формируются в процессе обучения и не меняются после. Карты активации же зависят от изображения, поданного в сеть на обработку.

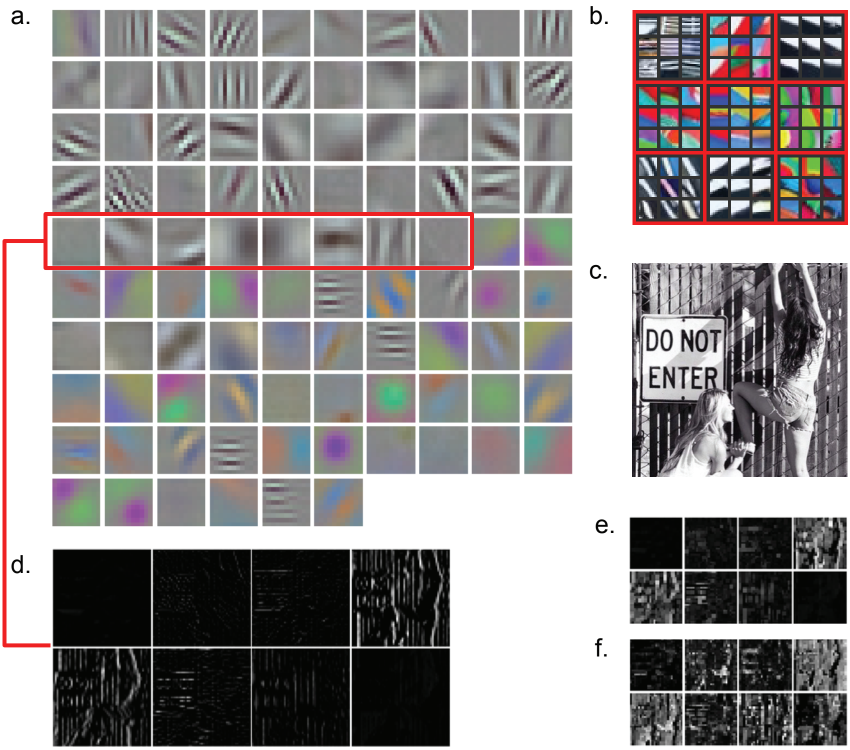


Рис. 3. Первый сверточный слой (conv1): *a*) фильтры слоя; *b*) фрагменты изображений, вызывающих максимальную активацию отдельного нейрона слоя (для девяти случайно выбранных нейронов представлены 9 лучших фрагментов); *c*) оригинальное изображение; *d*) карты активации выделенных фильтров; *e*) карты активации после осуществления операции Pooling; *f*) карты активации после проведения операции нормализации.

Фильтры первого сверточного слоя (Рис. 3) выделяют простые признаки вроде цвета и наклона линии — они универсальны практически для всех сетей, работающих с изображениями, вне зависимости от выполняемой задачи. Изображения, вызывающие максимальную активацию в нейронах данного слоя, представляют из себя фрагменты небольшого размера, содержащие линии, углы, решетки и т.д.

На втором сверточном слое (Рис. 4) фильтры не выглядят настолько же легко интерпретируемыми, размер рецептивного поля увеличивается и изображения, подобранные для визуализации максимальной активации, содержат менее тривиальные формы.

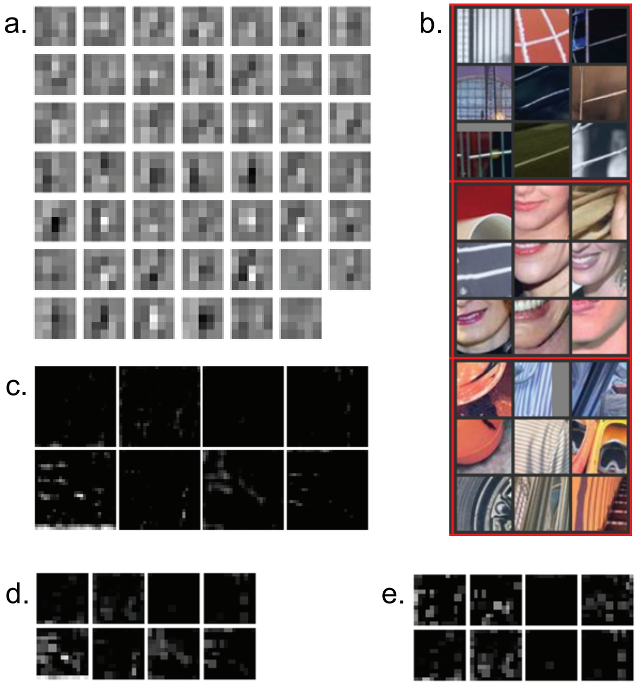


Рис. 4. Второй сверточный слой (conv2): *a*) фильтры слоя; *b*) фрагменты изображений, вызывающих максимальную активацию отдельного нейрона слоя (для трех случайно выбранных нейронов представлены 9 лучших фрагментов); *c*) карты активации выделенных фильтров. *d*) карты активации после осуществления операции Pooling; *e*) карты активации после проведения операции нормализации

Нейроны последующих сверточных слоев охватывают все большую часть изображения и реагируют на все более сложные формы объектов, вплоть до выделения различных категорий объектов. Тем не менее, в указанных слоях сохраняется зависимость реакции нейрона от пространственного расположения признака, то есть если рыжий кот в правом верхнем углу активирует определенный нейрон, то этот же самый кот в нижнем левом углу уже не вызовет в нем ответа. И только в полносвязных слоях сети достигается инвариантное — независимое от положения и размера — распознавание образов.

Каждый нейрон полносвязного слоя получает информацию от всех нейронов предыдущего слоя, что позволяет учитывать все возможные пространственные локации объектов. Для некоторых нейронов возможно определить признаки/объекты, выделяемые ими на изображениях, для других же это сделать довольно затруднительно (Рис. 5).



Рис. 5. Изображения, наилучшим образом активирующие три случайно выбранных нейрона первого полносвязного слоя (fc6). Для каждого нейрона подобраны девять изображений

Последний из полносвязных слоев содержит количество нейронов, равное количеству классов в исходной задаче классификации, т. е. в случае задачи детекции текста на изображении, этот будет состоять из двух нейронов: один, представляющий категорию «текст», другой — представляющий категорию «не текст» (Рис. 6).

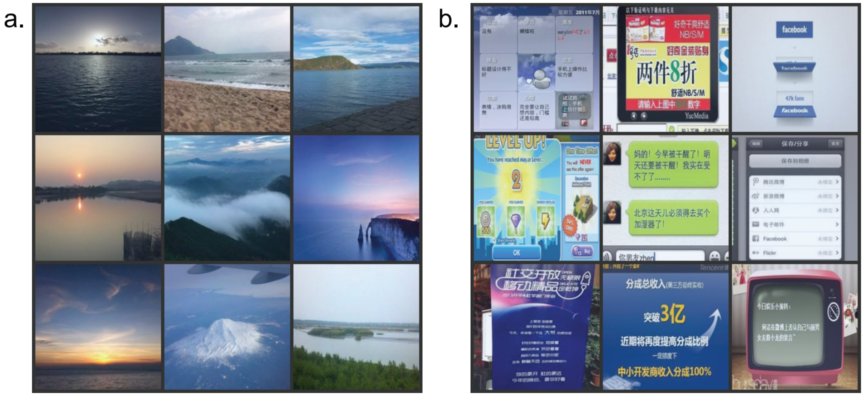


Рис. 6. Изображения, наилучшим образом активирующие нейроны последнего полносвязного слоя сети (fc8), обученной для задачи детекции лиц на изображениях. *a*) нейрон, отвечающий за категорию «не лица». *b*) нейрон, отвечающий за категорию «лица»

Вышеприведенные примеры визуализации созданы при помощи метода подбора изображения (или части изображения) из базы, который аналогичен методу построения кривой настройки, где из всех «предъявленных» нейрону изображений выбираются те, которые вызывают максимальный уровень активации. Качество подбора и интерпретации результатов в этом случае зависит от объема предъявляемой базы и наличия в ней подходящего изображения. Но это не единственный подход к поиску стимула вызывающего максимальную активацию. Другим методом является генерация изображения — на основе предоставленного изображения либо «с нуля» [Nguyen et.al., 2016]. Этот способ направлен на создание «идеального» стимула для каждого отдельного нейрона и не имеет аналогов в нейрофизиологических методах: случайно сгенерированное изображение итеративно изменяется на уровне отдельных пикселей с целью получения более сильного ответа выбранного нейрона. Качество полученного результата зависит от количества проведенных итераций, начальной инициализации входного стимула и выбранного способа генерации изображения. На Рис. 7 приведены примеры стимулов, вызывающих максимальную активацию отдельных нейронов в ИНС, полученные при помощи подбора лучшего изображения из базы (а) и генерации изображения (б).

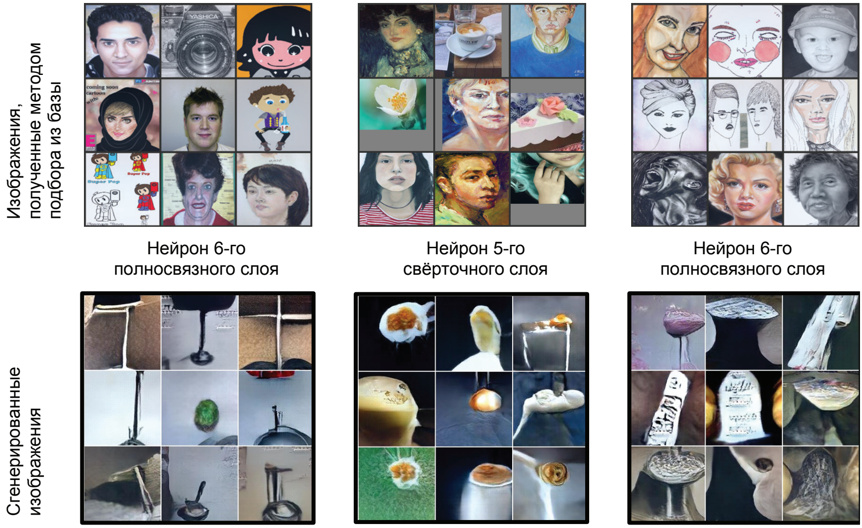


Рис. 7. Примеры стимулов, активирующих отдельные нейроны ИНС, полученные при помощи различных методов.

Как видно из Рис. 7, предположения, сделанные на основе результатов подбора изображений, могут быть смещены в сторону реальных объектов, имеющихся в базе. Например, все три нейрона в верхнем ряду можно отнести к нейронам, детектирующих лица. При рассмотрении результатов, полученных методом генерации, мы, во-первых, не встречаем лиц (что может объясняться неспособностью генератора создавать стимулы в виде лиц), а во вторых, видим набор абстрактных форм, который, так или иначе, объединен каким-то набором общих признаков (цвет, форма, текстура, расположение). Сопоставляя результаты обоих методов мы начинаем замечать сходства в подобранных и сгенерированных стимулах, что позволяет сделать более корректные предположения о функциях нейрона.

Заключение

Наблюдается ряд аналогий в процессе обработки зрительной информации искусственными и биологическими нейросетями. Визуализация внутренней структуры сетей подтверждает найденные сход-

ства: на нижних слоях, изображение пропускается через фильтры ориентации линий, цвета и частотности, которые сформировались естественно после просмотра искусственной сетью миллионов изображений, а на более высоких слоях образы становятся комплекснее и начинают обретать семантическое значение.

При этом, в исследованиях биологических сетей время записи активности нейрона и количество предъявлений ограничены уровнем физической нагрузки на испытуемых, необходимостью проведения нескольких сессий и сложностью сопоставления данных, полученных от разных участников эксперимента. Изучение представления информации в искусственных сетях, в свою очередь, позволяет зафиксировать состояние системы и проводить неограниченное количество экспериментов. Предложенные компьютерными науками методы визуализации функций отдельных нейронов могут быть полезны также и для изучения свойств биологических систем. Применение методов визуализации требует дополнительной адаптации технологии, но предварительные результаты позволяют говорить о важности взаимодополнения классических подходов с современными технологиями.

Литература

Cadieu, C. F., Hong, H., Yamins, D. L., Pinto, N., Ardila, D., Solomon, E. A., Majaj, N. J. & DiCarlo, J. J. (2014). Deep neural networks rival the representation of primate IT cortex for core visual object recognition. PLoS Comput Biol, 10(12), e1003963.

DiCarlo, James J., and David D. Cox. «Untangling invariant object recognition.» Trends in cognitive sciences 11.8 (2007): 333–341.

Fukushima, K., & Miyake, S. (1982). Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In Competition and cooperation in neural nets (pp. 267–285). Springer Berlin Heidelberg.

Güçlü, U., & van Gerven, M. A. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. Journal of Neuroscience, 35(27), 10005–10014.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097–1105).

LeCun, Y., Huang, F. J., & Bottou, L. (2004). Learning methods for generic object recognition with invariance to pose and lighting. In Computer Vision and Pattern Recognition, 2004. CVPR2004. Proceedings of the 2004 IEEE Computer Society Conference on (Vol. 2, pp. II-104). IEEE.

Manassi M., Sayim B., Herzog M. H. «When crowding of crowding leads to uncrowding» //Journal of Vision, 2013 .13 (13):10. doi: 10.1167/13.13.10.

Nguyen, A., Yosinski, J., Bengio, Y., Dosovitskiy, A., & Clune, J. (2016). Plug & play generative networks: Conditional iterative generation of images in latent space. arXiv preprint arXiv:1612.00005.

Sato, T., Uchida, G., Lescroart, M. D., Kitazono, J., Okada, M., & Tanifuji, M. (2013). Object representation in inferior temporal cortex is organized hierarchically in a mosaic-like structure. Journal of Neuroscience, 33(42), 16642–16656.

Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034.

Yosinski, J., Clune, J., Nguyen, A., Fuchs, T. & Lipson, H. (2015). Understanding neural networks through deep visualization. arXiv preprint arXiv:1506.06579.

Нейротехнологии диагностики и реабилитации

Глава 14.
Нейротехнологии управления целенаправленной деятельностью человека в условиях виртуальной среды

Е. Ю. Шелепин, С. В. Муравьева, Е. Г. Якимова, Ю. Е. Шелепин

Целенаправленная деятельность или рефлекс цели согласно представлениям И. П. Павлова является ключевым в поведении человека [Павлов, 1924]. Тренировка целенаправленной деятельности человека происходит с учителем и без учителя в течение всей жизни, она является основой воспитания человека, его духовных и моральных качеств, деловой, в том числе боевой подготовки. Тренировка целенаправленной деятельности должна быть основой и для восстановления, реабилитации человека после заболеваний мозга, приводящим к нарушениям в когнитивной и эмоциональной сфере. Цель проводимой работы дать обоснование и показать разработанные нами методы восстановления.

После работ И. П. Павлова, К. М. Быкова и В. Н. Черниговского по изучению кортико-висцеральных отношений, были проведены исследования Ю. Д. Кропотовым, Н. Н. Василевским, Н. В. Суворовым, Н. В. Черниговской и др. по тренировке с применением ЭЭГ биообратной связи в физиологии, медицине и в спорте [Василевский и др., 1974; Василевский 1988; Кропотов, 2010; Сороко, Трубачев, 2003; Черниговская, 1978, 1983], одновременно были развиты методы аутогенных тренировок [Лобзин, Решетников, 1986]. Применяемые восстановительные методы на основе биообратной связи и аутогенных тренировок, включающие в себя трудотерапию, занятия спортом и всеми видами искусств, в том числе музыкой и пением, определяют пути дальнейшего развития методов восстановительной психотерапии.

Погружение в виртуальную среду и клиника

Виртуальная среда — это мир, созданный с помощью цифровых технологий, который передается человеку через его ощущения: зре-

ние, слух. С каждым годом расширяется список областей применения виртуальной среды, охватывая не только традиционные области — компьютерные игры и кино, но и области образования, лечебной и восстановительной медицины (хирургии, лучевой диагностики) и военного сектора. Виртуальная и дополненная реальность является удобной средой, в которой может проходить и уже проходит разнообразная целенаправленная деятельность человека, даже столь сложная, как управление транспортными средствами (самолетом, автомобилем). В последние десятилетия виртуальную среду стали активно использовать для лечения и реабилитации пациентов в неврологических клиниках [Chawla, Foreman, 2011; Foreman et al., 2004; Korallo et al., 2004; Sandamas G., Foreman, 2015; Wilson, Foreman, 1993].

Найджел Фореман и его коллеги уже давно провели исследования точности оценки операторами пространственной информации при работе в виртуальной среде, по сравнению с реальной средой. Более того, они показали, как происходит передача полученных навыков ориентации в пространстве виртуальной среды в ориентацию в эквивалентной реальной среде. Они изучали этот перенос не только у здоровых испытуемых, но и у детей, пожилых испытуемых и у взрослых с физическими недостатками, которые плохо ориентируются в сложно организованном пространстве. Так было проведено обучение этих испытуемых в условиях виртуальной симуляции многоуровневого комплексного торгового центра и затем в условиях уже реального данного помещения. Была показана эффективность восстановления целенаправленной деятельности пациентов с помощью предварительного обучения в условиях виртуальной среды. Такое предварительное обучение облегчает дальнейшее поведение в сложной реальной среде [Stanton et al., 2000, 2002].

В последнее время в мировой и отечественной практике для восстановления мозговых нарушений пациентов, например, после инсультов, в комплексе с медикаментозными средствами используют восстановительные методы, базирующиеся на погружении пациента в виртуальную среду.

На основе этих технологических подходов, в лаборатории физиологии зрения Института физиологии им. И. П. Павлова была со-

здана нейротехнология с использованием интерактивной виртуальной реальности для реабилитации пациентов с нейрокогнитивными расстройствами в психиатрических и неврологических клиниках. Изобретение относится к области реабилитационной медицины, в частности к восстановительной терапии и неврологии, и может быть использовано в комплексной реабилитации пациентов с нейрокогнитивными расстройствами, вызванными психоневрологической патологией (шизофренией, депрессией, фобиями и т.д.) или черепно-мозговой травмой.

В настоящее время существует Британский прототип нашего комплекса. Так, для решения задач восстановительной терапии пациентов после инсультов с целью улучшения их состояния и качества жизни Э. Пелахом в Йорке были созданы стенды, обеспечивающие улучшение походки и зрительного контроля осанки и движений [Durgin, Pelah, 1999; Durgin, Pelah, 2005; Casey, Pelah, 2010; Lehtinen et al., 2009; Nanapragasam, Pelah, 2009; Pelah et al., 2002; Pelah, Thurrell, 2001; Pelah et al., 2015; Thurrell, Pelah, 2005]. Были использованы исследованные ранее Э. Пелахом эффекты зрительной обратной связи [Pelah, Barlow, 1996; Pelah, Barlow, 1996]. В 2011 году профессор Э. Пелах пригласил Евгения Шелепина для работы. Усовершенствованную систему тренировок они собрали во Флориде в Атлантическом университете. Дальнейшее развитие системы шло в Санкт-Петербурге в ИФ РАН, где в 2014 году и был создан тренажер нового поколения для восстановления психических функций [Шелепин и др., 2016].

Внешне тренажеры Института физиологии им. И.П. Павлова РАН напоминают применяемые в спортзалах (фитнес центрах) тренажеры: беговые дорожки и экран с динамическими изображениями. Система захвата изображения и соответствующее программное обеспечение позволяли спортсмену или пациенту контролировать собственные движения.

Из аналогов наиболее близкими по технической сущности к предлагаемому изобретению являются способ и устройство, раскрытые в патенте WO 0112269 A1 [Pyles, Stanziano, 2001]. Изобретение относится к модульному фитнес-устройству, состоящему из стационарного фитнес-устройства (велотренажер, беговая дорожка)

с установленным на нем видеодисплеем, электронным медиа-интерфейсом и датчиком физических упражнений. Интерфейс позволяет подключать дисплей к одному или нескольким различным электронным носителям, включая персональный компьютер, видео/аудио магнитофон, телевизионную приставку, игровую консоль и т. д. Предусмотрено подключение к Интернету. Пользовательские интерфейсы для управления устройством могут включать в себя комбинацию голосовой активации, сенсорного коврика для мыши или клавиатуры, подвижно прикрепленной к устройству [Pyles, Stanziano, 2001].

Имеется и Российский прототип**.** В России был разработан способ, основанный на использовании компьютерного медицинского мультимедийного комплекса, предложенный профессором, министром здравоохранения Российской Федерации В. И. Скворцовой с коллегами. Ими получен патент РФ № 48783 на полезную модель, содержащую регистрирующие датчики, компьютер с соответствующим программным обеспечением и шлем или очки виртуальной реальности [Скворцова и др., 2010]. Согласно способу В. И. Скворцовой и ее коллег физиологические параметры пациента регистрируются, как и параметры стимуляции и отображаются в результате их переработки в рабочей программе и на мониторе пациента, шлеме или очках виртуальной реальности в виде специальной диагностической, тренажерной, игровой или другой задачи, которой занимается пациент. Наблюдения изменения регистрируемых биомеханических, физиологических или иных параметров позволяют осуществлять как самоконтроль пациента, так и контроль пациента врачом. Метод очень близок к методу, предложенному Эдаром Пелахом, но вместо сигнала биологической обратной связи в виде пляшущего человечка «аватара», как это реализовано в системе Пелаха, пациент должен был видеть на экране дельфина в воде и отождествлять себя в виде дельфина. Другое важное отличие состоит в том, что погружение в виртуальную среду в методе В. И. Скворцовой и ее коллег происходит более глубоко, так как одеты очки (шлем). Эти отличия принципиальны. Метод профессора В. И. Скворцовой предназначен для начала восстановительной терапии при грубых нарушениях в поведении человека, в его состоянии после инсульта и даже для

применения в острый период. Врач наблюдает за выполнением рабочего процесса со своего монитора и при необходимости может вне сознательного контроля пациента изменить настройки рабочего процесса для достижения определенных значений физиологических параметров. Таким образом, авторы используют биологическую обратную связь (БОС) и виртуальную среду, для чего проводят установку очков и шлема на голову пациента, установку датчиков движения на голову, туловище и тазовую область пациента, загрузку программного обеспечения, состоящего из модели виртуальной среды и элементов управления, и проводят направленную тренировку координированных движений головы, туловища и тазовой области посредством среды виртуальной реальности и датчиков движения. Для управления движениями рук или ног пациента во время сеанса виртуальной игры на его пораженной конечности закрепляется специальный сенсор, содержащий в своем составе акселерометры, позволяющие получать информацию о перемещении конечности. Посредством данного устройства движения регистрируются и передаются в игровую виртуальную среду, которая по зрительному каналу воспринимается пациентом. Относительным недостатком способа и устройства является использование 3D-очков или шлема виртуальной реальности, которые у многих пациентов вызывают побочные явления: утомление глаз, головокружение, тошноту, в результате могут резко ухудшить состояние пациента.

Способ-прототип

Из приведенных аналогов наиболее близкими по технической сущности к предлагаемому изобретению являются способ и устройство, раскрытые в заявке WO 0112269 (A1), МПК A61B 5/00, опубликованной 22.02.2001. Способ-прототип основан на активном движении человека при погружении его в виртуальную среду и биологической обратной связи по зрительному каналу. Кроме того, в тренажер встроен компьютер с заложенными в нем программами виртуальной среды и подвижного объекта, цифровая видеокамера, отслеживающая движение человека. Отслеживаемый объект распознается системы компьютерного зрения [Haemaelaeinen, 2008]. Данные видеонаблюдения могут быть использованы для определе-

ния скорости движения человека и его идентификации. Достоинство тренажера, который по сути является фитнес оборудованием, по сравнению с другими аналогами того же назначения состоит в том, что он интерактивен. Указанное изобретение относится в целом к оздоровительному комплексу и может использоваться для укрепления сердечно-сосудистой и опорно-двигательной систем, повышения иммунитета и жизненного тонуса пациентов. Однако прототип не может быть использован в лечебных целях для целенаправленной реабилитации пациентов с психоневрологической патологией, так как в нем осуществляется контроль только за физическим состоянием человека и не отслеживается состояние его психической сферы, вследствие чего отсутствует возможность оценки хода восстановительной терапии когнитивных и эмоциональных функций центральной нервной системы. Прототип не предусматривает мотивационный характер поведения пациента из-за отсутствия постановки перед ним конкретной задачи и необходимости ее решения, что важно для пациентов с патологической доминантой.

Цель проведенной нами работы. Существующие методы очень близки к нашей разработке. Мы во многом базировались именно на них. Поэтому для восстановления целенаправленной деятельности при психических нарушениях (шизофрения, депрессия, пограничные состояния, хронический стресс, невроз навязчивых состояний или обсессивно-компульсивный синдром) в лаборатории физиологии зрения Института физиологии им. И. П. Павлова в результате развития технологий, был создан собственный тренажер для восстановления не только двигательных, но и психических, когнитивных и эмоциональных функций. Рассмотрим принцип действия данного тренажера, представляющего собой простое решение важной задачи.

Предложен способ реабилитации пациентов, сочетающий использование виртуальной среды, как целеобразующего фактора, и двигательных нагрузок (движущаяся беговая дорожка или велотренажер) синхронизированные по нагрузке с движением пациента в виртуальной среде. В виртуальной среде создают ситуации, требующие от пациента принятия решения по выходу из них с максимальной концентрацией внимания на достижении поставленной

цели. Одновременно от пациента требуют локальной концентрации внимания на виртуальном подвижном объекте. При этом на протяжении всего сеанса реабилитации осуществляют контроль состояния пациента по электрической активности мозга, миограмме, измерению частоты сердечных сокращений с возможностью обновления создаваемой ситуации по результатам проведенного контроля. Перед сеансом и после его окончания осуществляют контроль состояния зрительной системы пациента по измерению ее временных и пространственно-частотно-контрастных характеристик с возможностью подбора в виртуальной среде полосы пространственных и временных частот, обеспечивающих компенсацию нарушений контрастной чувствительности пациента. Способ позволяет создать нейротехнологии реабилитации пациентов с нейрокогнитивными расстройствами за счет перестройки работы зрительной системы и включения тех временных и пространственно — частотных каналов зрительного анализатора, которые обеспечивают стимуляцию адаптивных возможностей мозга, связанных с когнитивными функциями.

Внимание пациентов, традиционно переключаем с целей, связанных с болезнью, на другую цель, которую предлагаем в игровом сюжете в виртуальной среде, в соответствии с нашими представлениями о свойствах каналов зрительной системы, передающих осознаваемую и неосознаваемую информацию. Мы формируем структуру и смысл изображений, так чтобы спектр изображений виртуальной среды был согласован с видимой (осознаваемой и неосознаваемой) пространственно-частотной и временной полосой пропускания каналов зрительной системы. Полосу пропускания устанавливаем на основании измерения контрастной чувствительности пациента. Полученные данные позволяют передать информацию, корректирующую поведение пациента, по тому пространственно частотному каналу зрительной системы, по которому полученные сведения воспринимаются либо осознанно, либо неосознанно. Именно избирательная стимуляции низкочастотных или высокочастотных пространственно-частотных каналов, позволяет осуществить передачу по ним необходимых управляющих сигналов для «восстановления пациента».

Наш аппаратно-программный комплекс

Задача, на решение которой направлена наша работа, заключается в создании нейротехнологий реабилитации пациентов с нейрокогнитивными расстройствами за счет перестройки работы зрительной системы и включения тех временных и пространственно-частотных каналов зрительного анализатора, которые обеспечивают стимуляцию адаптивных возможностей мозга, связанных с когнитивными функциями. Для достижения этого результата используем целенаправленное активное поведение пациента в результате погружения пациента в зрительную виртуальную или дополненную среду. Пациент должен осуществлять реальное движение на беговой дорожке или стационарном велосипеде с регулируемой нагрузкой в соответствии с контекстом игры или видеоролика. Проводим контроль физиологических параметров, позволяющих осуществлять оценку состояния пациента. Внешне это напоминает тренировки на велотренажерах или беговых дорожках в фитнес центрах.

Новым является то, что в виртуальной среде создают ситуации, требующие от пациента принятия решений по выходу из них с концентрацией внимания на достижении поставленной игровой цели. Особенность игры включает ориентацию в пространстве, обнаружение, распознавание объектов и их счет. Объекты соответствовали определенным характеристикам, характерным для парво-системы (объектное зрение) или магно-системы (пространственное зрение). Помимо решения глобальной задачи в некоторых играх в задачу пациента входит локальная концентрация внимания на виртуальном подвижном объекте. По результатам проведенного контроля производится обновление игровой ситуации. Перед сеансом и после его окончания проводится контроль состояния зрительной системы пациента по измерению временных и пространственно-частотно — контрастных характеристик зрительной системы, с возможностью подбора в виртуальной среде полосы пространственных и временных частот, обеспечивающих стимуляцию.

Для реализации данного способа создан аппаратно-программный восстановительный комплекс. Аппаратные средства включают управляющий компьютер, с дополнительным панорамным дисплеем (большим жидкокристаллическим монитором или проекцион-

ным экраном), установленный перед пациентом. Данный комплекс, включает средства, обеспечивающее активное движение — стационарный велотренажер или беговую дорожку. Имеются датчики контроля изменения нагрузки тренажера. На протяжении всего сеанса реабилитации осуществляется контроль физического, когнитивного и эмоционального состояния пациента — текущей электрической активности мозга (ЭЭГ), электрокардиограммы (ЭКГ), контроль движения глаз, мимики. Режим нагрузок устанавливается и изменяется в зависимости от физического состояния человека, стимулирующим компьютером, синхронизированным с управляющим компьютером, обеспечивающим систему контроля физиологических параметров, характеризующих общее физическое и эмоциональное состояние пациента, состояние его зрительной системы. Программное обеспечение включают динамические изображения виртуальной среды — игры, интерактивные коммуникации, позволяет осуществлять подбор управляющих сигналов. Подбор определяется задачей восстановления и количественными данными о состоянии мозга пациента, что обеспечивает перестройку работы нейронных сетей, участвующих в зрительном восприятии, принятии решения и прогнозировании двигательной активности, а также стимуляцию адаптивных возможностей мозга, связанных с когнитивными функциями.

Программа виртуальной среды, представляет собой интерактивный видеоряд, имитирующий, например, велопрогулку по различным ландшафтам с различным рельефом местности, полностью синхронизированный с велотренажером. Программа позволяет проецировать на экран монитора схематизированное изображение пациента, помещенного в виртуальную среду с возможностью перемещения в ответ на движения реального объекта. Эта возможность развивает способ Эдара Пелаха и отдаленно напоминает дельфина В. И. Скворцовой и соавторов. Наш способ тренирует не только собственные движения на беговой дорожке или велотренажере, а формирует целенаправленную деятельность. По своей сути комплекс представляет собой игровой автомат, в котором игрок взаимодействует с изображением, отображенном на плоском экране, максимально концентрируя внимание на изображении и цели. Но при этом комплекс снабжен устройствами контроля, позволяющими оценить

физическое и психическое состояние «игрока» и внести в «игру» соответствующие коррективы. Сущность предложенного способа состоит в создании технологий реабилитации пациентов с нейрокогнитивными расстройствами. Это достигается с одной стороны за счет перестройки работы нейронных сетей и включения тех временных и пространственно-частотных каналов зрительного анализатора, которые в заданных пороговых условиях обеспечивают неосознаваемый уровень обучения. С другой стороны — за счет переключения со сформированных болезнью целей на цели, задаваемые и формируемые виртуальной игровой средой, в результате чего происходит «переобучение» пациента. Переключению на достижение цели помогают физические нагрузки, выполняющие роль подкрепления при выработке новых временных связей в процессе обучения достижению цели, выработки рефлекса цели.

Развиваемая в настоящее время часть нашей работы — компьютерный контроль переориентации пациента осуществляют искусственные нейронные сети глубокого обучения на основании информации от ЭЭГ, мимики лица и контроля движений глаз (см. главу Р. О. Малашина в данной книге).

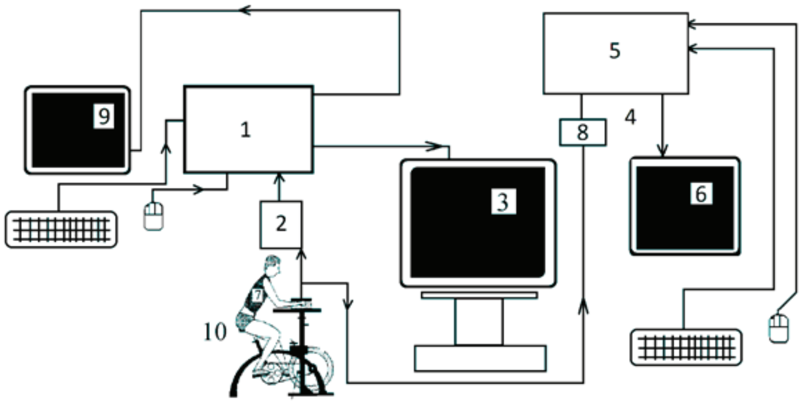


Рис. 1. Схема способа реабилитации с использованием зрительной биообратной связи пациентов погруженных в виртуальную среду [Шелепин и др., 2016, Патент R U2654767 C1]. Цифры, нанесенные на блоки, раскрыты в тексте

Последовательность процедуры реабилитации включает следующие этапы: размещение пациента на кресло велотренажера пред экраном монитора; загрузку программного обеспечения, состоящего из виртуальной среды и элементов контроля; объяснение задачи пациенту и проведения им пробных движений; проведение сеанса реабилитации. На рис. 1 и 2 представлены схема и фотографии одной из комбинаций аппаратно-программного комплекса реабилитации виртуальной средой, которая содержит компьютер 1 с программами изображения виртуальной среды (VRE) и подвижного объекта в виде изображения пациента, видеокамеру 2, обеспечивающую видеонаблюдение за движениями пациента, например, выполненную с возможностью преобразования видеосигнала в цифровой сигнал, монитор 3, представляющий собой мультимедийное средство, обращенное в сторону пациента, систему регистрации физиологических и функциональных параметров 4, включающую в себя компьютер 5, монитор 6, датчик регистрации физиологи-



Рис. 2. Внешний вид аппаратно-программного комплекса для реабилитации виртуальной средой в лаборатории физиологии зрения Института физиологии им. И. П. Павлова РАН [Шелепин и др., 2016, Патент R U2654767 C1].

ческих параметров 7 (частота сердечных сокращений), состояния ЦНС 8 (электроэнцефалограф с датчиками), состояния зрительной системы 9 (айтрекер); и средство, обеспечивающее активное движение пациента 10 (велотренажер или беговая дорожка). Комплекс предназначен для восстановления зрительных когнитивных нарушений, вызванных психоневрологической патологией на основе использования программ с отслеживанием движения пациента в виртуальной среде.

Необходимо отметить, что в лаборатории создано программное обеспечение, позволяющее воспроизводить движение человека в виртуальном пространстве и создавать виртуальный аватар собственного тела пациента (рис. 3). Программа использует данные сенсора Kinect для построения 3D-модели человека и воспроизводит ее на экране в реальном времени. В то же время данные о трех-

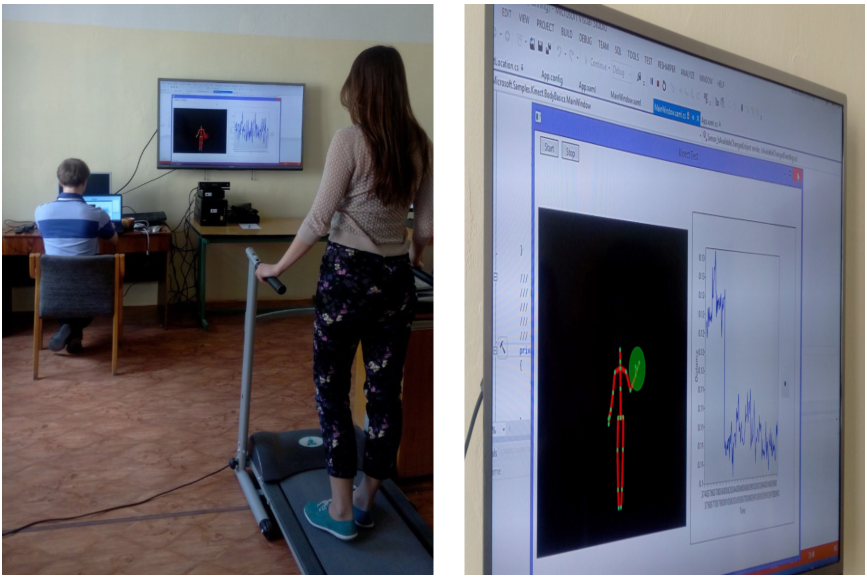


Рис. 3. Работа аппаратно-программного комплекса для реабилитации пациентов после двигательных нарушений в результате центральных нарушений виртуальной средой в режиме захвата движений и построения аватара тела пациента. Аватар служит для коррекции движений в качестве зрительной обратной связи. Система является развитием систем проф Эдара Пелаха из Йорка (Великобритания)

мерных координатах каждого контролируемого сенсором сустава записываются в базу данных. Этот аватар отображается на экране, в то время как пациент находится на беговой дорожке. Система получает различную информацию о параметрах ходьбы и тела пациента (например, длина и количество шагов, симметричность позы тела, темп ходьбы и т.д.). В дальнейшем эта информация используется врачом (исследователем) для контроля за состоянием пациента, фиксирования прогресса в его реабилитации. На экран также могут выводиться задания, которые должен выполнять пациент. На основе этих данных выводится статистика по исследуемым величинам. На основе данных манипуляций у человека формируется обратная связь о положении своего тела. Если у человека есть проблемы в моторной сфере, то они будут осознаваться человеком, соответственно, он будет стараться их исправить.

Курс тренировки на аппаратно-программном комплексе реабилитации виртуальной средой обеспечивал одновременное воздействие на целостную деятельность организма человека.



Рис. 4. Система реабилитации пациентов с нарушениями в когнитивной и эмоциональной сфере психической деятельности в виртуальной среде со зрительной обратной связью

Тренажер включает в себя велотренажер, обеспечивающий с одной стороны более полное погружение в виртуальную среду, с другой — наряду с физическими нагрузками, улучшающими кровообращение, способствует подкреплению цели за счет перестройки обмена биологически активных веществ и увеличения их выброса в кровь [Шелепин и др., 2016, Патент R U2654767 C1].

Результаты реабилитационных тренировок показали, что в процессе воздействия виртуальной среды с двигательной нагрузкой происходит активное формирование у пациента «рефлекса цели», формирование целенаправленной деятельности с подкреплением в виде сопряженной мышечной активности, направленной на достижение цели (пользуясь Павловской терминологией, «подкрепление выработки рефлекса цели»). Этот принцип является крайне важным в эффективной реабилитации пациентов, так как в результате достижения цели происходит перестройка работы всего организма, мышечная активность улучшает состояние всех систем организма. Человеку становится приятно от предлагаемых ему процедур (данный метод), улучшается «качество жизни» пациента (см. главу С. В. Муравьевой и Ю. Е. Шелепина).

Сущность предложенного способа состоит в создании нейротехнологий реабилитации пациентов с нейрокогнитивными расстройствами. Это достигается с одной стороны за счет перестройки работы зрительной системы и включения тех временных и пространственно-частотных каналов зрительного анализатора, которые в заданных пороговых условиях обеспечивают неосознаваемый уровень обучения. Контроль переориентации пациента осуществляют искусственные нейронные сети глубокого обучения на основании информации от ЭЭГ, мимики лица и контроля движений глаз (см. главу Р. О. Малашина).

Болезни, вызванные психоневрологической патологией, представляют собой полиморфные заболевания, нарушающие процесс мышления, поведения и эмоциональную сферу человека (см. статью С. В. Муравьевой и Ю. Е. Шелепина в данной книге). Согласно многочисленным нейрофизиологическим и психофизическим исследованиям выявлено, что одними из основных особенностей таких заболеваний являются сенсорные нарушения, среди которых особое место занимают нарушения зрительной системы, играющей ключевую роль для адаптации человека в окружающей среде. В свою очередь, нарушения работы зрительной системы приводят к сбою когнитивных функций: нарушению восприятия и идентификации изображений, нарушению процесса сложной визуальной группировки объектов, нарушению хранения в памяти зрительной информации

и ее использования при распознавании, возникновению неадекватных эмоциональных реакций на подаваемые зрительные сигналы и т.д.

Были проведены исследования на здоровых испытуемых и пациентах с шизофренией по тестированию разработанного комплекса. Проведено исследование характера изменений когнитивных зрительных вызванных потенциалов (ЗВП), контрастной чувствительности и фМРТ данных паттернов мозговой активации у пациентов с шизофренией с продолжительностью заболевания от 1 до 5-х лет до воздействия виртуальной средой с двигательной нагрузкой и после воздействия. В результате реабилитационных тренировок пациентов с различными диагнозами (шизофренией, депрессивными и фобическими расстройствами) было показано, что заявляемый метод позволяет вовлечь в активную работу магно- (низкочастотные) и парво- (высокочастотные) каналы зрительной системы, что характеризуется выраженным увеличением контрастной чувствительности системы, увеличением амплитуды когнитивных компонентов зрительных вызванных потенциалов и появлением новых активных зон обработки зрительной информации (по электрофизиологическим данным и данным фМРТ исследований). Это в свою очередь позволяет приблизить состояние пациентов с нейрокогнитивными расстройствами к параметрам здоровой группы испытуемых.

Заключение

Оптимизация физических и семантических характеристик виртуальной среды за счет сигналов обратной связи, регулирующих деятельность человека, позволяют разработать систему рекомендаций для создания новых технологий виртуальной среды, оптимальной для деятельности человека в разные возрастные периоды. Известно, что с возрастом ухудшается распознавание и ориентация в пространстве. Поэтому разработка интерактивной виртуальной среды на основе зрительной биологической обратной связи между параметрами этой среды и физиологическими показателями деятельности человека является актуальным направлением исследований. Предлагаемый метод является с нашей точки зрения продол-

жением работ с биологической обратной связью, традиционным направлением в отечественной физиологии. В данном случае объектом регуляции мозга является сам мозг, как важнейший адаптивный орган организма человека. Главное это то, что восстанавливается целенаправленная деятельность, ключевая в поведении человека [Павлов, 1924].

Литература

Быков К. М. Кора головного мозга и внутренние органы, М., 1942.

Василевский Н. Н., Суворов Н. Б., Трубачев В. В. Адаптивная модуляция временной структуры эндогенных ритмов нейронных популяций мозга // В сб.: Системный анализ интегративной деятельности нейрона /. М.: Наука, 1974.С. 115–124.

Василевский Н. Н. Биоуправление с обратной связью // Биоуправление: теория и практика. Новосибирск, 1988. С. 17–42.

Кропотов Ю. Д. Количественная ЭЭГ, когнитивные вызванные потенциалы мозга человека и нейротерапия. 2010.

Лобзин В. С., Решетников М. И. Аутогенная тренировка (справочное пособие для врачей). Л.: Медицина. 1986. — 280 с.

Никишина И. С., Пономарев В. А., Кропотов Ю. Д. Бета 1-биоуправление при лечении синдрома нарушения внимания с гиперактивностью: динамика изменений в количественной ЭЭГ после курса лечения. Нейроиммунология 2004; 2: 2: 77.

Скворцова В. И., Иванова Г. Е., Скворцов Д. В. Способ реабилитации больных в острой стадии инсульта с использованием биологической обратной связи и виртуальной реальности. Патент РФ, RU2 432 971 C1. Приоритет 02.04.2010.

Сороко С. И., Трубачев В. В. Нейрофизиологические и психофизиологические основы адаптивного биоуправления. СПб.: Политехника-сервис, 2010. — 607 с.

Павлов И. П. Двадцатилетний опыт объективного изучения выс-

шей нервной деятельности (поведения) животных. М.-Пг.: Госиздат, 1923. Гл. «Рефлекс цели» [М.: Наука, 1973. С. 214–218].

Черниговская Н.В. Адаптивное биоуправление в неврологии,—JL: Медицина, 1978, С. 49–76.

Черниговская Н.В., Мовсисянц С.А., Тимофеева А.Н. Клиническое значение адаптивного биоуправления.—JL: Медицина, 1983.—С. 14–59.

Черниговский В.Н. Интероцепторы, М., 1960.

Черниговский В.Н. Нейрофизиологический анализ кортико-висцеральной рефлекторной дуги, Л., 1967.

Шелепин Ю.Е., Муравьева С.В., Шелепин Е.Ю., Якимова Е.Г., Шелепин К.Ю. Способ реабилитации пациентов с использованием виртуальной среды //патент RU (11) 2654767 (13) C1 Дата начала отсчета срока действия патента: 02.12.2016.

Allen L.E., Slater M.E., Proffitt R.V., Quarton E., Pelah A. A new perimeter using the preferential looking response to assess peripheral visual fields in young and developmentally delayed children. Journal of Aapos: the Official Publication of the American Association For Pediatric Ophthalmology and Strabismus / American Association For Pediatric Ophthalmology and Strabismus. 16: 261–5. 2012 PMID22681943 DOI: 10.1016/j.jaapos.2012.01.006.

Budzynski T.H. From EEG to neurofeedback. Introduction to quantitative EEG and Neurofeedback. Eds.: Academic Press 1999; 65–79.

Casey R,. Pelah A., Cameron J., Lasenby J. Symposium On Applied Perception in Graphics and Visualization. 2010: 160. DOI: 10.1145/1836248.1836282.

Chawla S., Foreman N. Forms of Interaction in Virtual Space: Applications to Psychotherapy and Counselling., DOI: 10.1007/978–3–642–21657–2_33, Source DBLP, Conference: Universal Access in Human-Computer Interaction. 6th International Conference, UAHCI 366,

2011, Orlando, FL, USA, 2011, Proceedings, Part IV.

Durgin F. H., Pelah A. Visuomotor adaptation without vision? Experimental Brain Research. 127: 12–8. 1999 PMID10424410 DOI: 10.1007/s002210050769.

Haemaelaeinen P. Object tracking in computer vision; WO2008046963 (A1); 2008–04–24. Also published as: EP2080168 (A1); FI121981 (B); FI20060926 (A); US2010322472 (A1).

Foreman N.,. Sandamas G., and Newson D. Distance underestimation in virtual space is sensitive to gender but not activity-passivity or mode of interaction.// CyberPsychology & Behavior 7(4):451–7, September 2004. DOI:10.1089/cpb.2004.7.451. https://www.researchgate.net/publication/8381345_Distance_Underes-timation_in_Virtual_Space_Is_Sensitive_to_Gender_But_Not_Activity-Passivity_or_Mode_of_Interaction [accessed Jul. 29 2018].

Korallo L., Foreman N., Boyd S., Moar M., Coulson M. Can multiple «spatial» virtual timelines convey the relatedness of chronological knowledge across parallel domains? Computers & Education 58(2): 856–862, February 2012, DOI: 10.1016/j.compedu.2011.10.011.

Nanapragasam A., Pelah A., Cameron J., Lasenby J. Visualizations for locomotor learning with real time feedback in VR Proceedings — Apgv 2009: Symposium On Applied Perception in Graphics and Visualization. 139. DOI: 10.1145/1620993.1621033.

Pelah A., Barlow H. B. Visual illusion from running. Nature. 381: 283. 1996 // PMID8692265 DOI: 10.1038/381283a0.

Pelah A., Barlow H. B. An illusion of accelerated self-motion following treadmill jogging Investigative // Ophthalmology and Visual Science. 37: S524, 1996.

Pelah A, Thurrell A. E. Reduction of perceived visual speed during locomotion: Evidence for quadrupedal perceptual pathways in human? Journal of Vision. 1: 307a. DOI: 10.1167/1.3.307, 2001.

Pelah A, Thurrell A. E, Berry M. Reduction of perceived visual speed during walking: Evidence against the involvement of attentional or vestibular mechanisms Journal of Vision. 2: 630a. DOI: 10.1167/2.7.630, 2002.

Pelah A., Barbur J., Thurrell A, Hock H. S. The coupling of vision with locomotion in cortical blindness. Vision Research. 110: 286–94. 2015 PMID24832646 DOI: 10.1016/j.visres.2014.04.015.

Pyles Nathan; Stanziano Steven M. Interactive fitness equipment, Applicant(s): [US] EPIX INC; WO0112269 (A1) — 2001–02–22; Classification A63B21/00; A63B24/00; Priority number(s): US19990374324 19990813. Also published as: WO0112269 (A9), AU7139700 (A), US2002055418 (A1).

Sandamas G., Foreman N. Active Versus Passive Acquisition of Spatial Knowledge While Controlling a Vehicle in a Virtual Urban Space in Drivers and Non-Drivers // SAGE Open, July-September 2015: 1–9, DOI: 10.1177/2158244015595443.

Wilson P., Foreman N. Transfer of information from virtual to real space: Implications for people with physical disability. Eurographics Technical Report Series. ISSN.1017–4656, pp.21–25. (1993).

Глава 15.
Восстановление нарушений целенаправленной деятельности у пациентов с психоневрологической патологией путем погружения в интерактивную виртуальную среду

С. В. Муравьева, Ю. Е. Шелепин

В последние годы все большее внимание уделяется изучению нейрофизиологических механизмов, лежащих в основе когнитивных и эмоциональных нарушений у пациентов с психоневрологической патологией. Среди них особое место занимают рассеянный склероз, шизофрения, депрессии и фобии. Актуальность исследований обусловлена тем, что этими заболеваниями в первую очередь страдают люди трудоспособного, чаще молодого возраста. Частыми проявлениями этих заболеваний являются нарушения работы зрительной системы, играющей ключевую роль для адаптации человека в окружающей среде. Возможности измерения отклонений в работе зрительной системы достигают квантовых порогов.

Зрительная система здорового человека состоит из множества каналов с определенными пространственно-частотными характеристиками, которые обеспечивают описание окружающего мира [Campbell, Robson, 1968]. Две крайние формы (морфологически отдельные и функционально различные пути) — низкочастотные каналы (магно-каналы), которые обеспечивают «пространственное зрение» (описание грубых пространственных признаков, важных для ориентации в пространстве), высокочастотные (парво-каналы) — «объектное зрение» (цвет, детали формы, контуры, размер объекта и т. д.).

Для некоторых неврологических и психических заболеваний характерны нарушения работы этих систем, в частности, при рассеянном склерозе [Муравьева и др., 2008, 2013, 2015] и шизофрении [Шошина, 2012, 2013; Муравьева, 2015; Муравьева и др., 2017; Muravyova at al., 2016; Murav`eva et al., 2016].

Объединяющими эти заболевания факторами являются схожие аутоиммунные и генетические механизмы их возникновения

[Stevens, Schizophrenia Bulletin, 1988]. Для пациентов с преобладаниями нарушений в работе парво-системы характерны жалобы на быстрое цветовое утомление, на нарушение цветового зрения, восприятия деталей формы и размера объекта. Для пациентов с преобладаниями нарушений в работе магно-системы характерны: диплопия, нарушение ориентации в пространстве, нарушения ощущения положения объекта в пространстве, нарушение восприятия направления движения объекта.

Раньше в области физиологии зрения применялся принцип «пассивного восприятия»: неподвижно сидящему перед экраном испытуемому предъявлялись статичные или движущиеся визуальные стимулы, и измерялась реакция на них. Но в естественных условиях восприятие — это активный процесс, при котором субъект должен активно взаимодействовать с окружающей средой: иметь возможность перемещаться относительно наблюдаемых предметов [Гибсон, 1988].

В настоящее время благодаря появлению систем «виртуальной реальности» стало возможным более точное моделирование естественных условий восприятия визуальных сцен. При этом возможности, предоставляемые современными системами синтеза и обработки изображений, позволяют создавать визуальные среды с любыми требуемыми пространственно-частотными характеристиками — как близкими к реальным, так и отличными от них, предназначенными для избирательного воздействия на различные подсистемы зрительного анализатора. Использование систем «виртуальной реальности» открывает широкую возможность для медицины — разработки технологий для компенсации утраченных человеком функций [Кузнецов и др., 2014; Schroeder, 2008; Wilson, Foreman, 1993].

На основе этого в лаборатории физиологии зрения Института физиологии была создана нейротехнология с использованием интерактивной виртуальной реальности для реабилитации пациентов с нейрокогнитивными расстройствами в психиатрических и неврологических клиниках. Задача — создание нейротехнологии реабилитации пациентов за счет перестройки работы зрительной системы и включения тех временных и пространственно-частотных каналов зрительного анализатора, которые страдают у этих пациентов. Они

связаны с когнитивными функциями и адаптацией. Комплекс был применен для реабилитации пациентов с шизофренией с продолжительностью заболевания от 0 до 10 лет.

Шизофрения — полиморфное психическое заболевание, нарушающее процесс мышления, поведения и эмоциональную сферу человека. Основной проблемой при шизофрении является дезадаптация человека в окружающей среде. Существенную часть публикаций, посвященных этой теме, составляют исследования связи нарушений адаптации с особенностями когнитивной сферы, и попытками разработки методов восстановления когнитивных функций [Dias et al., 2011; Симонова и др., 2014].

Известно, что для оценки состояния магно- и парво-систем достаточно чувствительными являются методы зрительных вызванных потенциалов и контрастной чувствительности.

В последнее время в мировой и отечественной практике для восстановления мозговых нарушений пациентов в комплексе с медикаментозными средствами используют терапевтические методы, базирующиеся на погружении пациента в виртуальную среду.

Для контроля эффективности реабилитации пациентов был выбран электрофизиологический метод — метод измерения когнитивных зрительных вызванных потенциалов, т.к. он является наиболее чувствительным для оценки изменений зрительной системы и когнитивных функций у пациентов, а также позволяет контролировать эффективность медикаментозной терапии [Plomp, 2013; Муравьева и др., 2015, 2017].

У здоровых наблюдателей было показано, что с помощью метода когнитивных зрительных вызванных потенциалов, зарегистрированных при выполнении тестов по классификации изображений, можно выделить компоненты, которые предположительно отражают процессы обработки зрительной информации на разных ее этапах [Моисеенко и др., 2015]. Эти же исследования были проведены и на пациентах с психическими расстройствами. У этих пациентов были выявлены характерные изменения амплитуды когнитивных компонентов зрительных вызванных потенциалов [Муравьева и др., 2017].

В исследовании принимали участие 36 пациентов с параноидной формой шизофрении (F20.0 по МКБ-10) с продолжительностью

заболевания от 1 до 5 лет в возрасте от 21 до 34 (20 мужчин и 16 женщин), а также контрольная группа — 26 здоровых испытуемых в возрасте от 18 до 30 лет (14 мужчин и 12 женщин). У всех наблюдателей острота зрения была не менее 0,9, рефракция соответствовала норме.

Условия проведения исследований соответствовали этическим нормам Хельсинкской декларации всемирной медицинской ассоциации. Все пациенты и здоровые испытуемые подписывали информированное согласие.

Среди симптомов, которые наблюдались у большинства пациентов, которые участвовали в исследовании, необходимо отметить эмоциональную однообразность, гипомимичность, вместе с тем напряженность и настороженность, тревожность. Среди типичных симптомов шизофрении в анамнезе всех пациентов отмечалась амбивалентность (двойственное, противоречивое отношение субъекта к объекту, характеризующееся одновременной направленностью на один и тот же объект противоположных импульсов, возникающих внезапно и независимо от обстоятельств), симптомы «соскальзывания». Были характерны также бредовые идеи преследования, бредовые идеи воздействия, слуховые галлюцинации (голоса). Все пациенты, которые принимали участие в исследованиях, находились в стационаре на лекарственной терапии нейролептиками третьего поколения. Для пациентов было характерно доминирование продуктивных симптомов над негативными. У пациентов не было расстройств мышления и дезорганизации поведения.

Исследование состояло из трех этапов. Первый этап включал в себя первичный контроль состояния зрительной системы до воздействия виртуальной средой: измерение контрастной чувствительности и исследование когнитивных зрительных вызванных потенциалов.

Для измерения пороговой контрастной чувствительности использовали программу «Эрготест 2.0», позволяющую измерять пороговые контрасты синусоидальных решеток с заданными пространственными частотами, предъявляемых на экране монитора. Для этого пациента сажали в кресло перед экраном монитора. Решетки предъявляли в пространственно-частотном диапазоне от 0.25 до

17 цикл/угл. град (0,25; 0,56; 1,12; 2,25; 4,50; 9,00 и 17,00 цикл/угл. град) с постепенным нарастанием контраста от 0 до 1. Расстояние между монитором и глазами испытуемого составляло 1,5 м. Пациенту было предложено нажимать на кнопку мыши, как только заметит появление решетки и отпускать, как только решетка исчезнет.

После этого проводилось ЭЭГ — исследование и измерение когнитивных зрительных вызванных потенциалов, полученных методом усреднения на стимулы, обработанные вейвлетной фильтрацией в низкочастотном и высокочастотном диапазоне пространственных частот (Рис.1).

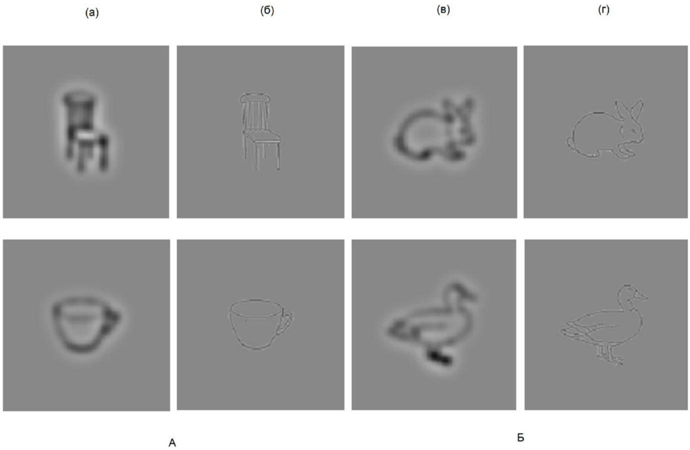


Рис. 1. Примеры 2-х категорий черно-белых изображений — объектов неживой (А) и живой природы (Б), отфильтрованных с помощью вейвлетного преобразования в области низких (а), (в) и высоких пространственных частот (б), (г).

Данные ЭЭГ регистрировались во время выполнения тестов по классификации изображений объектов живой и неживой природы. Параметры фильтров высокой и низкой частоты составляют 0,53 Гц и 30 Гц, также используется режекторный фильтр — 45–55 Гц. Регистрация электроэнцефалограммы (ЭЭГ) производилась с помощью 31-канального компьютерного электроэнцефалографа «Мицар — 202» с частотой дискретизации 250 Гц и программы «WinEEG». На голову пациента одевалась электроэнцефалографическая шапочка с 31 электродами, расположенными на поверхности головы в соот-

ветствии с модифицированной системой 10–20 в отведениях Fp1; Fpz; Fp2; F7; F3; Fz; F4; F8; T3; C3; Cz; C4; T4; T5; P3; Pz; P4; T6; O1; O2. Референтные электроды располагались на мочках ушей, а заземляющий — в лобной области.

Порог статистической достоверности результатов регистрации вызванных потенциалов рассчитывался с помощью критерия Вилкоксона в программе Statistica и составил 5%, $p < 0,05$. Статистическую обработку данных измерения пороговой контрастной чувствительности производили методом двухвыборочного Т-теста с неравными дисперсиями (достоверным отличием считалось значение $p<0,05$).

На этапе первичного исследования (до воздействия виртуальной среды) у пациентов с шизофренией были выявлены следующие достоверные изменения. При предъявлении изображений, содержащих высокочастотную часть спектра, у пациентов было зарегистрировано достоверное снижение амплитуды компонентов P100, N170 и P250 в затылочной области (в отведениях O1 и O2), по сравнению со здоровыми испытуемыми (Рис. 2). Наблюдалось также достоверное снижение амплитуды компонентов N100 и N250 и при отведении из центральной и лобной области (в отведениях C3, Cz, C4; F3, Fz, F4), амплитуды компонента N170 — в отведениях T5 и T6. Снижение амплитуды компонента P500 в теменной и центральной области в отведениях P3, Pz, P4; C3, Cz, C4. От содержания изображений (живые или не живые объекты) этот эффект не зависел.

При сравнении соотношения амплитуд компонентов зрительных вызванных потенциалов на стимулы, отфильтрованные по пространственным частотам у контрольной группы и пациентов, наблюдается достоверное отличие компонентов P100 (N100) в затылочных, центральных и лобных отведениях: амплитуда на стимулы низких частот достоверно выше, чем на высоких (рис. 3А, 3Б). А амплитуда компонента P170 у контрольной группы в теменных (P3, Pz, P4), в центральном отведении Cz и лобных отведениях (F3, Fz, F4) достоверно не отличается, в центральных отведениях (C3, C4) — на высоких частотах выше, чем на низких (рис. 3А, 3Б). А у пациентов наблюдается обратная картина: амплитуда этого компонента достоверно выше на низких пространственных частотах, чем на высоких. У контрольной группы амплитуда компонента N250 (P250) в заты-

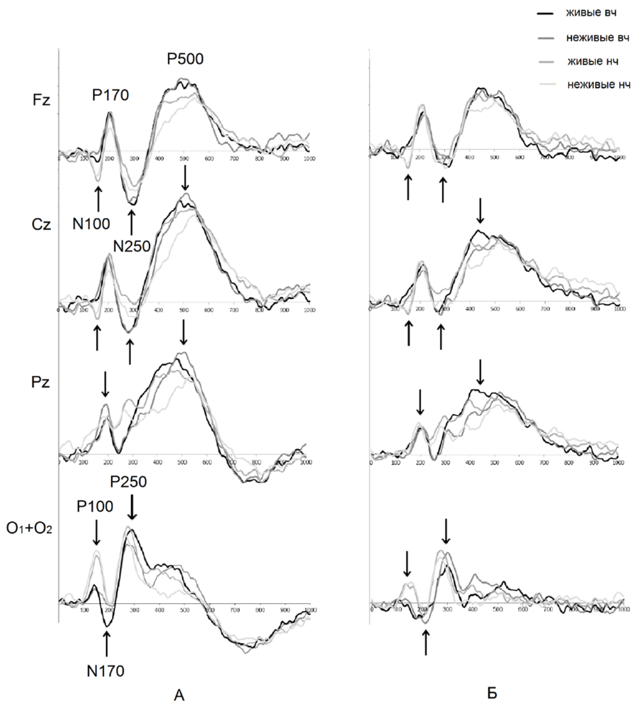
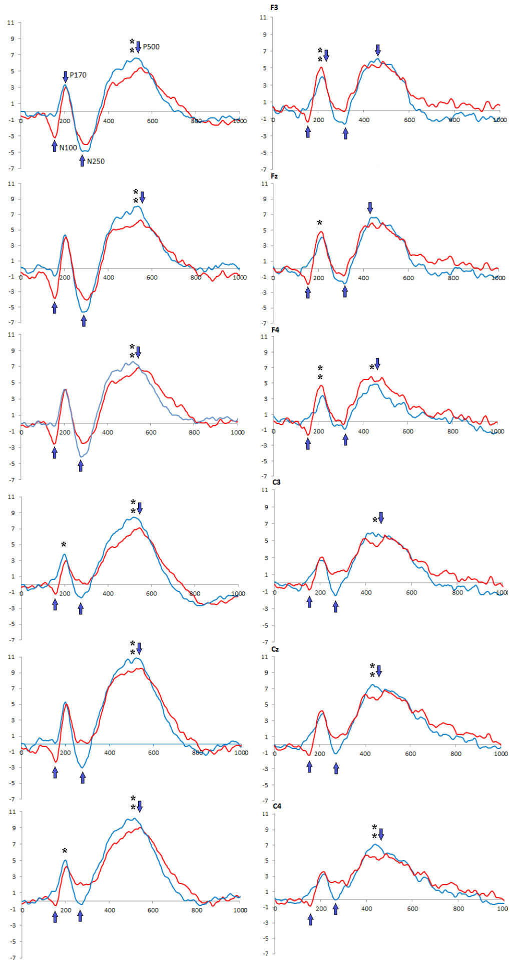


Рис. 2. Усредненные вызванные потенциалы во время выполнения инструкции по различению семантических признаков в изображениях у контрольной группы и пациентов с шизофрении в центральном лобном отведении (Fz), центральном отведении (Cz), центральном теменном отведении (Pz), в затылочных отведениях (усреднение по двум затылочным отведениям (O1 и O2)).

А. Усредненные вызванные потенциалы контрольной группы. Б. Усредненные вызванные потенциалы пациентов с шизофренией. Стрелками показаны достоверные различия ($p <$ 0,05) по семантическим и пространственно-частотным характеристикам изображений

лочных, центральных и лобных отведениях на стимулы высоких частот достоверно выше, чем низких, а у пациентов достоверно не отличается. Амплитуда компонента P500 в теменных (P3, Pz, P4), центральных (C3, Cz, C4) и лобных отведениях (F3, Fz, F4) на стимулы высоких частот достоверно выше, чем низких. От содержания изображений (живые или не живые объекты) этот эффект не зависел. Все эти изменения показаны на рис. 3 на примере усредненных вызванных потенциалов на стимулы живых объектов.



А Б

Рис. 3. Усредненные вызванные потенциалы во время выполнения инструкции по различению семантических признаков в изображениях у контрольной группы и пациентов с шизофрении в центральном лобном отведении (Fz), центральном отведении (Cz), центральном теменном отведении (Pz), в затылочных отведениях (усреднение по двум затылочным отведениям (O1 и O2)).

А. Усредненные вызванные потенциалы контрольной группы. Б. Усредненные вызванные потенциалы пациентов с шизофренией.Стрелками показаны достоверные различия ($p < 0{,}05$) по семантическим и пространственно-частотным характеристикам изображений.

При предъявлении изображений, содержащих низкочастотную часть спектра, было зарегистрировано достоверное снижение амплитуды компонентов N170 (P170) при предъявлении изображений, различающихся по семантическим признакам в затылочных и теменных отведениях у пациентов по сравнению с контрольной группой. Достоверное снижение амплитуды компонентов P250 (N250) наблюдалось в затылочных и лобных отведениях, P500 — в затылочных, теменных и центральных отведениях. При сравнении соотношения амплитуд внутри каждой группы для компонентов N170 (P170) — и у пациентов, и у контрольной группы. P250 (N250) — только у контрольной группы. Эти изменения были зарегистрированы в затылочной, теменной, центральной и лобной областях.

Измерение пороговой контрастной чувствительности выявило достоверное снижение в области высоких пространственных частот у пациентов по сравнению с данными контрольной группы.

Второй этап — непосредственно нейрореабилитация — курс сенсомоторной нагрузки. Курс состоял из десяти сеансов. Продолжительность сеанса 15 минут. Во время сеанса (тренинга) пациент размещался перед панорамным экраном большого углового размера (46.7 угл.градусов, расстояние до глаз — 140 см) на кресле велотренажера. В задачу пациента входило внимательно смотреть на экран, на котором предъявлялась виртуальная программа, осуществлять движение педалями и отслеживать движение объектов, появляющихся на экране и задаваемых пациенту в начале тренинга. Предъявлялись виртуальные среды — видеоряд, имитирующий велопрогулку от первого лица по различным ландшафтам с различным рельефом местности. Уровень сопротивления менялся в зависимости от рельефа. В задачу пациента входило не только внимательно смотреть на ландшафт, но выполнять когнитивную задачу: акцентировать внимание на отдельных его элементах для того, чтобы производить, например, счет людей, машин или однотипных дорожных знаков, появляющихся на экране. Стимулы соответствовали определенным характеристикам, характерным для парво-системы (объектное зрение) или магно-системы (пространственное зрение). Объекты различались по цветам (черные, белые или цветные), размеру, расположению, могли быть стационарными или находиться в движе-

нии. Курс обеспечивал одновременное воздействие и на зрительную (стимулировалось как объектное, так и пространственное зрение), на сердечно-сосудистую и на опорно-двигательную систему. После курса были проведены электрофизиологические и психофизические исследования такие, как и до курса.

Третий этап — контрольные измерения когнитивных зрительных вызванных потенциалов и исследования контрастной чувствительности после курса сенсомоторной нагрузки. Стимулы и условия проведения те же, что и до курса.

На рис. 4 показаны данные зрительных вызванных потенциалов до курса и после него на стимулы живых объектов высоких и низких пространственных частот на примере одного пациента. После курса воздействия заметно увеличение амплитуды компонентов P100 в затылочной области, P170 в центральных (C3, Cz, C4) и лобных (F3, Fz, F4) областях на низких и высоких частотах; N250 в теменных областях и P500 в центральных и лобных областях на высоких частотах. Необходимо отметить, что меняется и соотношение амплитуд: появляется различие амплитуды компонента P170 в теменной (Pz) и центральной (Cz) области и P350 — P500 в теменной, центральной и задней лобной области.

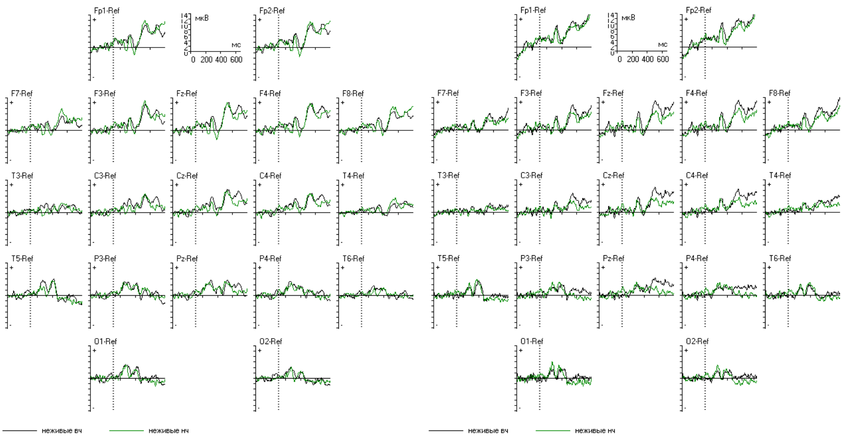


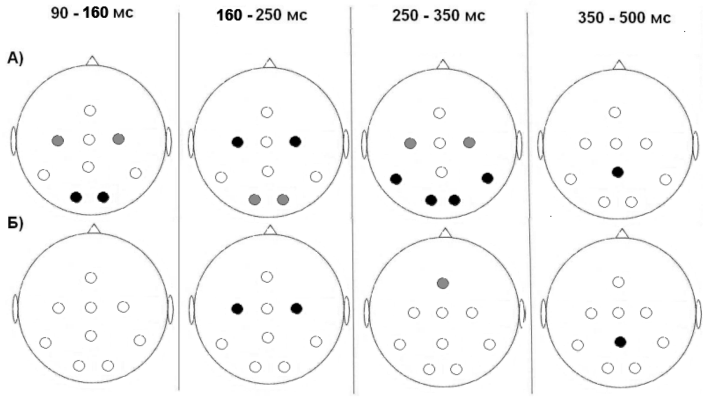
Рис. 4. Зрительные вызванные потенциалы на стимулы живых объектов высоких и низких пространственных на примере данных одного пациента до курса воздействия и после него

После статистической обработки данных группы пациентов были получены следующие результаты. В зависимости от пространственно-частотных признаков (А) изображения при повторном исследовании (после воздействия виртуальной средой) по сравнению с данными до воздействия было зарегистрировано достоверное увеличение амплитуд компонентов N170 (P170) в затылочной (O1 и O2) и центральной области (Cz); P250 (N250) в затылочной (O1 и O2) и теменной (Pz) области и P350—P500 в лобной (F3, Fz, F4), центральной (Cz, C3, C4) и теменной (P3, Pz, P4) области при предъявлении изображений живых и неживых предметов, отфильтрованных по высоким пространственным частотам, появляется достоверное отличие по сравнению с низкими пространственными частотами. (Рис. 5А).

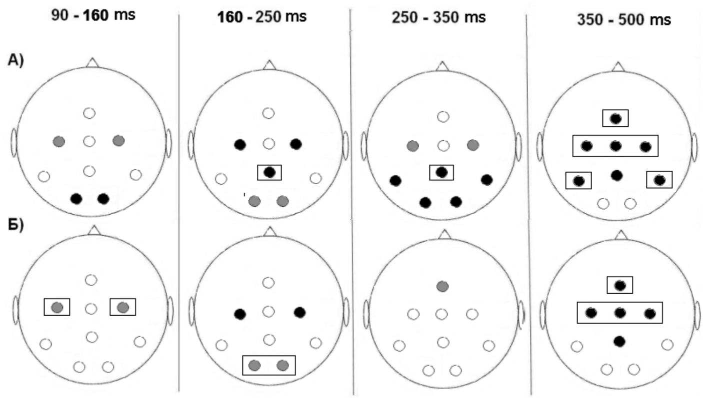
При сравнении изображений по семантическим признакам (Б) наблюдалось достоверное увеличение амплитуды компонента N100 в центральной области (C3, C4) при предъявлении изображений, отфильтрованных по низким пространственным частотам, компонента N170 в затылочной области (O1 и O2) и P350—P500 в лобной (F3, Fz, F4) и центральной (Cz, C3, C4) при предъявлении изображений, отфильтрованных по высоким пространственным частотам, появляется достоверное отличие по сравнению с низкими пространственными частотами (Рис. 5Б).

Таким образом, после курса сенсо-моторной нагрузки с интерактивной виртуальной средой у пациентов наблюдается активация теменной, центральной и затылочной областей как по пространственно-частотным, так и по семантическим признакам: возникновение новых зон дифференцировки зрительных образов. Наиболее выраженные изменения характерны для поздних компонентов P350—P500 (Рис. 5Б).

Современные литературные данные о том, какая именно система, магно- или парво-, поражается при шизофрении, противоречивы. Большинство работ посвящено тому, что при этом заболевании поражаются изолированно магно-каналы [Kim, Park, 2011; Oribe et al., 2013; Skottun, Skoyles, 2013], реже встречаются данные о нарушении работы парво-каналов [Doniger et al., 2002; Ardekani et al., 2003; Lalor et al., 2012] или двух систем одновременно [Lalor et al., 2012].



5А



5Б

● области мозга, где были выявлены достоверные различия (p<0,05) в амплитуде вызванных потенциалов, имеющих отрицательную полярность

● области мозга, где были выявлены достоверные различия (p<0,05) в амплитуде вызванных потенциалов, имеющих положительную полярность

○ области мозга, где не было выявлено статистически достоверных различий

Рис. 5. Сравнительный анализ амплитуды компонентов когнитивных зрительных вызванных потенциалов пациентов до и после воздействия виртуальной средой.

Инструкция испытуемым — различение семантических признаков в изображениях.

(*а*) — зависимость амплитуды от пространственной частоты изображений, (*б*) — зависимость амплитуды от семантических свойств изображений. Рамкой выделены области, в которых появляются достоверные различия в отношении амплитуд компонентов после проведения курса воздействий

Необходимо отметить, что Донигер с соавторами в своей работе описывает результаты электрофизиологических исследований, демонстрировавших снижение амплитуды компонентов зрительных вызванных потенциалов, генерируемых вентральной областью зрительной системы и полученных при использовании стимулов различных контрастов у пациентов с шизофренией ранних стадий [Doniger et al., 2002]. А в работе Ардекани с соавторами описаны данные, полученные методом функциональной магнитно-резонансной томографии с использованием «низкочастотных» и «высокочастотных» визуальных стимулов у пациентов с шизофренией с первым эпизодом, свидетельствующие о снижении восприятия только высоких пространственных частот (парво-система), тогда как у хронических больных наблюдается нарушение восприятия всего спектра частот [Ardekani et al., 2003].

На основании полученных нами результатов исследования когнитивных зрительных вызванных потенциалов можно предположить, что у пациентов с параноидной формой шизофрении продолжительностью заболевания от 1 года до 10 лет наблюдается доминирующая дисфункция парво-каналов в зрительной системе. Это подтверждают также результаты измерений пороговой контрастной чувствительности, демонстрирующие ее понижение в области высоких пространственных частот.

Исследования, проведенные для контроля изолированного воздействия курса физической нагрузки на зрительную систему человека до и после курса тренировок на велотренажере без применения виртуальной среды достоверных изменений у контрольной группы не выявили.

В результате исследований после сенсомоторной нагрузки с интерактивной виртуальной средой было выявлено, что заявляемый метод позволяет вовлечь в активную работу в первую очередь парво-(высокочастотные) каналы зрительной системы (именно те каналы, которые в большей степени страдают у этих пациентов), в меньшей степени — магно- (низкочастотные) каналы. Это характеризуется выраженным увеличением амплитуды когнитивных компонентов зрительных вызванных потенциалов и появлением новых активных зон дифференцировки зрительной информации. Это в свою очередь позволяет прибли-

зить состояние зрительной системы пациентов с нейрокогнитивными расстройствами к параметрам здоровой группы испытуемых. Также были получены данные об активации обработки зрительной информации как на ранних, так и на поздних этапах. Наиболее выраженные изменения были выявлены для поздних компонентов. Именно тогда, когда наступает самооценка принятого решения.

Для того, чтобы исключить предположение об эффекте обучения самому тесту при повторном использовании одного и того же теста на группе здоровых испытуемых было проведено двухразовое исследование с перерывом в 10 дней. Таким образом, были соблюдены условия исследования со стимуляцией интерактивной виртуальной средой. У этой группы после повторного исследования достоверного увеличения амплитуд компонентов вызванных потенциалов получено не было, таким образом эффект обучения не подтвердился.

Таким образом, можно сделать вывод, что использование интерактивной виртуальной среды у пациентов может быть основой для разработки технологий для компенсации утраченных нарушений зрительной системы у пациентов с психоневрологической патологией [Муравьева и др., 2015, 2017; Муравьева, 2017; Murav`eva et al., 2016].

Предлагаемый метод восстановления когнитивных функций может способствовать повышению эффективности лечения пациентов наряду с медикаментозными методами. Это становится возможным с одной стороны за счет восстановления работы зрительной системы и ее когнитивных функций. С другой стороны — за счет переключения с целей, связанных с болезнью и погружения в виртуальную среду, которое помогает отвлечься от патологической доминанты. А вовлечение в физические нагрузки обеспечивает улучшение периферического и мозгового кровообращения и всей деятельности организма в целом и активизирует выброс в кровь эндорфинов.

Список литературы:

Гибсон Дж. Экологический подход к зрительному восприятию, М. // Прогресс, 1988.

Кузнецов П. П., Чеботаев К. Ю., Узденов Б. И. Медицина и виртуальная реальность 21 века: Создание синтетических

сред, тренды, инновации // Врач и информационные технологии. 2014, № 3, С. 72–80

Моисеенко Г. А., Шелепин Ю. Е., Хараузов А. К., Пронин С. В., Чихман В. Н., Вахрамеева О. А. Классификация и распознавание изображений живой и неживой природы // Оптический журнал, 82, 10, 2015, С. 53–64.

Муравьева С. В., Дешкович А. А., Шелепин Ю. Е. Магно и парво-системы человека и избирательные нарушения их работы // Рос. физиол. журн. им. И. М. Сеченова. 2008. Т. 94. № 6. С. 637.

Муравьева С. В., Фокин В. А., Ефимцев А. Ю., Шелепин Ю. Е. Пространственно-частотные каналы зрительной системы при рассеянном склерозе // Сенс. Сист. 2013. Т. 27. № 2. С. 130.

Муравьева С. В., Бисага Г. Н., Пронин С. В., Брякилева Т. В., Шелепин Ю. Е., Влияние продолжительности заболевания на изменения зрительных вызванных потенциалов контрастной чувствительности у пациентов с рассеянным склерозом // Журн. высш. нервн. деятел. 2015. Т. 65. № 6. С. 711.

Муравьева С. В., Пронина М. В., Моисеенко Г. А., Шелепин Ю. Е. Исследование когнитивных нарушений при шизофрении на ранних стадиях заболевания // Материалы конференции. Всероссийская конференция с международным участием, посвященная 90-летию со дня основания Института физиологии им. И. П. Павлова РАН — СПб. 2015. С. 143–144.

Муравьева С. В., Пронина М. В., Моисеенко Г. А., Пневская А. Н., Поляков Ю. И., Кропотов Ю. Д., Пронин С. В., Шелепин Е. Ю., Шелепин Ю. Е. Исследование зрительных когнитивных нарушений при шизофрении на ранних стадиях заболевания и их коррекция при помощи интерактивных виртуальных сред // Физиология человека. 2017. Том 43, № 6. С. 24–36.

Муравьева С. В. Влияние виртуальной среды с малыми и большими изображениями на магно- и парво-клеточные пути зрительной системы пациентов с шизофренией // Всероссийская

Конференция с Международным участием «Клиническая нейрофизиология и нейрореабилитация». Санкт-Петербург. Специальный выпуск журнала «Вестник клинической нейрофизиологии». 2017. С. 82–83.

Симонова Н. А., Гарах Ж. В., Зайцева Ю. С., Шмуклер А. Б. Нейрофизиологические механизмы нарушения зрительного восприятия при шизофрении // Социальная и клиническая психиатрия. 2014. Т. 24, № 1. С.—81.

Шошина И. И., Шелепин Ю. Е. Контрастная чувствительность у больных шизофренией с разной длительностью заболевания // Рос. физиол. журн. им. И. М. Сеченова. 2013. Т. 99. № 8. С. 657.

Шошина И. И., Шелепин Ю. Е., Конкина С. А., Пронин С. В., Бендера А. П. Исследование парвоцеллюлярных и магноцеллюлярных зрительных каналов в норме и при психопатологии // Рос. физиол. журн. им. И. М. Сеченова. 2012. Т. 98. № 5. С. 657.

Ardekani B. A., Nierenberg J., Hoptman M. J., Javitt DC, Lim KO. MRI study of white matter diffusion anisotropy in schizophrenia // Neuroreport. 2003. Vol. 14. P. 2025.

Campbell F. W., Robson J. G. Application of Fourier Analyses to the Visibility of Gratings // J. Physiol. 1968. Vol. 197. P. 551.

Dias E. C., Butler P. D., Hoptman M. J., Javitt D. C. Early sensory contributions to contextual encoding deficits in schizophrenia // Arch. Gen. Psychiatry. 2011. Vol. 67, № 7. P. 654.

Doniger G. M., Foxe J. J., Murray M. M., Higgins B. A., Javitt D. C. Impaired visual object recognition and dorsal/ventral stream interaction in schizophrenia // Arch. Gen. Psychiatry. 2002. Vol. 59. P. 1011.

Kim D., Park S. Visual perception deficits associated with the magnocellular pathway in schizophrenia // Korean. Schizophr. Res. 2011. Vol. 14. P. 61.

Lalor E. C., De Sanctis P., Krakowski M. I., Foxe J. J. Visual sensory processing deficits in schizophrenia: Is there anything to the magnocellular account? // Schizophr. Res. 2012. Vol. 139. P. 24.

Muravyova S., Moiseenko G., Pronina M., Shelepin E., Shelepin Y. Dysfunction of parvo–systems and its stimulation in patients with schizophrenia with early stage of the disease // Perception. V 45, Supplement, 39th European Conference on Visual Perception, 2016. P. 14.

Murav'ova (Murav`eva) S., Moiseenko G. A., Pronina M., Pronin S. V., Shelepin E. Y., Shelepin Y. E. The impact of spatial-frequency filtering to recognize of stimuli in healthy subyects and patients with schizophrenia (psychophysical and electrophysiological methods of investigation) // Materials of the IEEE International Symposium «Video and Audio Signal Processing in the Context of Neurotechnologies», June 30 — July 2 2016 St. Petersburg.

Oribe N., Hirano Y., Kanba S., del Re E. C., Seidman L. J., Mesholam-Gately R., Spencer K. M., Mc Carley R. W., Niznikiewicz M. A. Early and late stages of visual processing in individuals in prodromal state and first episode schizophrenia: An ERP study // Schizophr. Res. 2013. Vol. 146. P. 95.

Plomp G., Roinishvili M., Chkonia E., **Kapanadze G., Kereselidze M., Brand A., Herzog M. H**. Electrophysiological evidence for ventral stream deficits in schizophrenia patients // Schizophr. Bull. 2013. Vol. 39, N3. P. 547.

Schroeder R. Defining Virtual Worlds and Virtual Environments // J. of Virtual Worlds Research. 2008. Vol. 1.№ 1. P. 1.

Skottun B. C., Skoyles J. R. Visually evoked potentials, NMDA receptors and the magnocellular system in schizophrenia // Acta Neuropsychiatrica. 2012. Vol. 24. P. 50.

Stevens Janice R. Schizophrenia and Multiple Sclerosis, Schizophrenia Bulletin, 1988, Volume 14. Issue 2. pp. 231–241

Wilson P., Foreman N. Transfer of information from virtual to real space: Implications for people with physical disability. Eurographics Technical Report Series. 1993. ISSN.1017–4656. pp. 21–25.

Глава 16.
Новая технология объективного измерения механизмов фильтрации и принятия решений

Г.А. Моисеенко, С.В. Пронин, Ю.Е. Шелепин

Целенаправленную деятельность человека обеспечивают механизмы зрительного восприятия, основные из них осуществляют определение глобальных и локальных информативных признаков объектов и сцены [Шелепин и др., 1985; 2009]. Нейрофизиологические основы этих механизмов — пространственно-частотные каналы с настройкой на разные пространственные частоты открытые Фергюсом Кемпбеллом и Джоном Робсоном [Глезер, 1993; Шелепин и др., 1985], согласованные фильтры согласно Н.Н. Красильникову, механизмы принятия решений и программирования действий [Красильников, Шелепин, 1997; Кропотов, 2010]. Управление вниманием позволяет в результате многоканальной пространственно-частотной фильтрации тестовых изображений усилить восприятие высоких пространственных частот, передающих информацию о локальных свойствах наблюдаемых изображений, или, наоборот, их подавить и выделить низкочастотную составляющую в изображении, дающую глобальное восприятие сцены. Роль внимания в переключении с одного канала на другой, в повседневной жизни, в боевой обстановке, в изобразительном искусстве мы обсуждали ранее [Шелепин и др., 1985]. Всегда при наблюдении естественных сцен происходит переключение с одного канала на другой с канала выделяющего малую деталь в сцене или сцену целиком. Причем в зависимости от задачи для нас, например, лицо может быть малой частью, деталью, крупной жанровой или батальной сцены, а в другой ситуации лицо это целая сцена и на ней разыгрывается спектакль, с помощью мимических мышц и положения взора. Переключения с локальных на глобальные признаки сцены обеспечивает механизм избирательного внимания. Чем выше пространственная частота, тем больше информации мы можем передать и воспринимать. Самая высокая воспринимая пространственная частота определяет нашу остроту зрения.

Обеспечивает остроту зрения деятельность самых высокочастотных каналов зрительной системы.

В процессе повседневного офтальмологического измерения остроты зрения пациента врач управляет вниманием пациента с помощью словесной инструкции. Принятие решения при выполнении пациентом даже простой задачи — определения ориентации решетки или положения разрыва в кольце Ландольта требует вовлечения структур головного мозга от первичных каскадов собственно зрительной системы до лобных областей коры и требует включения в работу механизмов принятия решений, планирования действий и обеспечения моторного ответа. Конечно, по сравнению с возможностями человека это довольно простая задача классификации, но и она требует активного участия пациента и добросовестного выполнения инструкции. Всегда важна задача объективизации клинических измерений. Существующая объективизация измерений остроты зрения или контрастной чувствительности с помощью регистрации вызванных потенциалов на решетки разной пространственной частоты позволяет оценить приход сигнала в зрительную кору, но не его осознание. Нейроофтальмологическая диагностика центральных поражений головного мозга, механизмов принятия решений о наблюдаемом изображении требует более изощренных методов. Существует целая система объективизации результатов измерений остроты зрения [Ковальская и др., 2013; Коскин и др., 2007, 2008].

Целью данного исследования является разработка технологии измерения характеристик зрительного восприятия, принятия решения и на этой основе оценка остроты зрения в режиме распознавания. В качестве техники измерения мы применили метод вызванных потенциалов.

Для усиления локальных признаков изображений применили высокочастотную фильтрацию, для их подавления и сохранения только гештальта изображения применили низкочастотную фильтрацию. На рис. 1 представлены примеры двух классов изображений.

Изображения предъявляли бинокулярно на электронно-лучевом мониторе с кадровой частотой 100 Гц и разрешением экрана 1024 × 768. Средняя яркость и контраст всех стимулов были одинаковы.

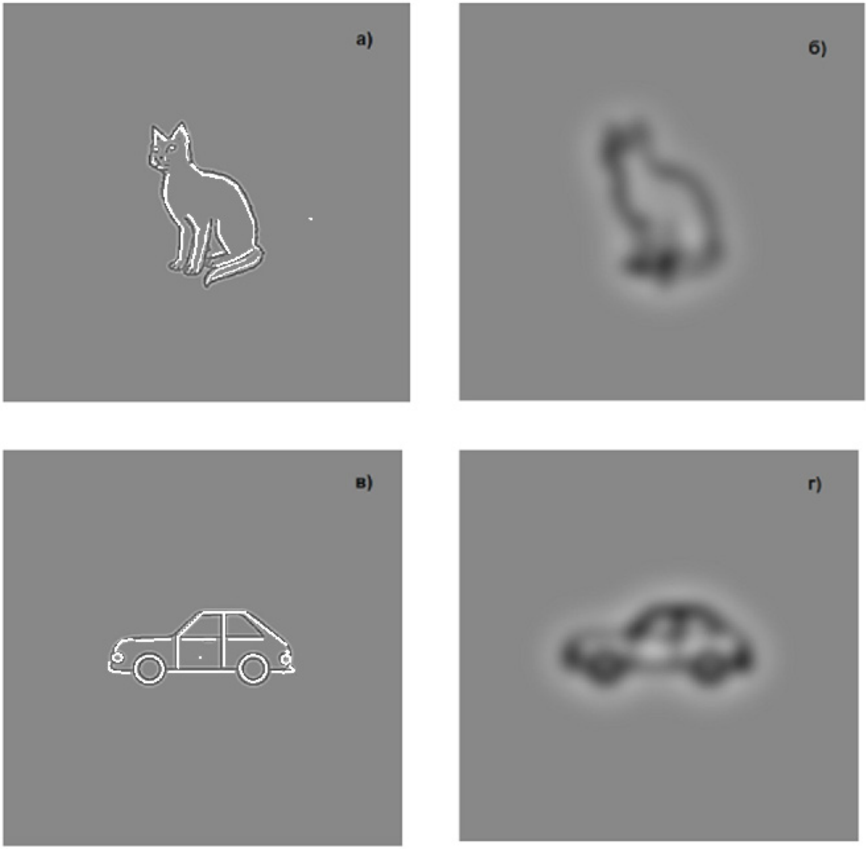


Рис. 1. Примеры двух классов черно-белых контурных изображений — объектов живой и не живой природы, отфильтрованных с помощью вейвлетной фильтрации (четкие-размытые)

Расстояние между монитором и глазами испытуемого составляло 1,5 м. Размер изображений — 3 угл. град. Изображения предъявляли в случайном порядке на 100 мс с интервалом в 1 с. Исследование состояло из двух серий с одинаковым дизайном эксперимента, но двумя разными инструкциями наблюдателю: 1) классифицировать изображения по признакам «живой/неживой»; 2) классифицировать изображения по признакам «четкий/размытый». В обеих сериях исследований применяли метод вызванных потенциалов. Регистрация

вызванных потенциалов проводили по схеме 10–20 с референтными ушными электродами. Регистрацию и обработку электроэнцефалограммы (ЭЭГ) проводили с помощью энцефалографа фирмы «Мицар-ЭЭГ-201». Для оценки статистической значимости различий амплитуды вызванных потенциалов использовали Т-критерий Вилкоксона (Wilcoxon matched pairs test). Результаты считали значимыми при $p < 0,05$. Для анализа вызванных потенциалов использовали пространственно-временной анализ. Анализировали основные компоненты вызванных потенциалов: компоненты P100 и N100, P170 и N170, P250 и N250, P300–500. Затем сравнивали результаты проведенных нами исследований между собой. В первом исследовании добровольно участвовал 21 испытуемый. Из них было 17 женщин и 4 мужчин в возрасте от 20 до 38 лет. Во втором исследовании 21 испытуемый, но из них было 13 женщин и 8 мужчин были в возрасте от 18 до 36 лет. Острота зрения испытуемых по таблицам Головина-Сивцева была не менее 1,0 или откорректирована до нормы зрения очковыми линзами. Все испытуемые были праворукими и без неврологических патологий.

Полученные результаты в исследованиях, проведенных с обеими инструкциями, мы можем трактовать как отражение параллельной обработки наблюдаемого сигнала по разным семантическим и физическим признакам. В исследованиях с применением 2-й инструкции (размыт объект или нет) были выделены изменения компонента N170 затылочных и височных отведений и компонента P200 при лобных отведениях на изменение семантики изображений. Т.е. различия по семантике в лобных областях происходили всегда, даже в том случае, когда задача испытуемого заключалась в классификации по физическим признакам изображений объектов, и он выполнял эту работу успешно. Это означает, что испытуемый одновременно с осознанным выбором и правильным нажатием соответствующей клавиши размытый «да» / «нет» осуществлял и нетребуемую от него назовем ее образно «неосознаваемую классификацию» на живые и неживые объекты. Испытуемый видел предъявляемые объекты, хотя его внимание инструкцией к действию привлекала другая задача — оценка степени размытия. Результаты особенностей важнейших компонентов вызванных потенциалов представлены на рис. 2.

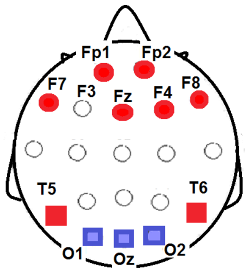


Рис. 2. Локализация положения электродов при регистрации вызванных потенциалов мозга. Синими квадратами отмечены положения электродов, от которых компоненты (N170) вызванных потенциалов различаются по амплитуде как на физические (размытый/четкий), так и на семантические (живой/неживой) характеристики стимулов. Красным цветом отмечены электроды, вызванные потенциалы от которых отличаются только по семантике (квадраты — компоненты N170, а круги — компоненты P200).

На основании полученных результатов можно сделать вывод о том, что, когда согласно инструкции испытуемый классифицирует наблюдаемые изображение объекта как «размытые/не размытые», исследователь может определить отличия по различиям амплитуд компонентов вызванных потенциалов в затылочной, височной и лобной коре. Это означает, что согласно инструкции испытуемый концентрирует внимание на осуществление классификации объектов размытый /четкий, но в мозге происходит классификация по критерию объект живой /неживой природы, так как человек осознанно или неосознанно, но видит эти объекты, даже в том случае если обращает или не обращает внимание на них.

Это наше наблюдение дает в руки врача — исследователя инструмент объективной оценки того что видит испытуемый, и как он проводит классификацию. Различная пространственно-частотная фильтрация изображения или управление дистанцией наблюдения дают возможность определить разрешающую способность наблюдателя независимо от принимаемого наблюдателем решения. Статистический контроль нажатий кнопки ответа позволяет определить стратегию поведения, избранную испытуемым. Пример такого анализа дан на рис. 3.

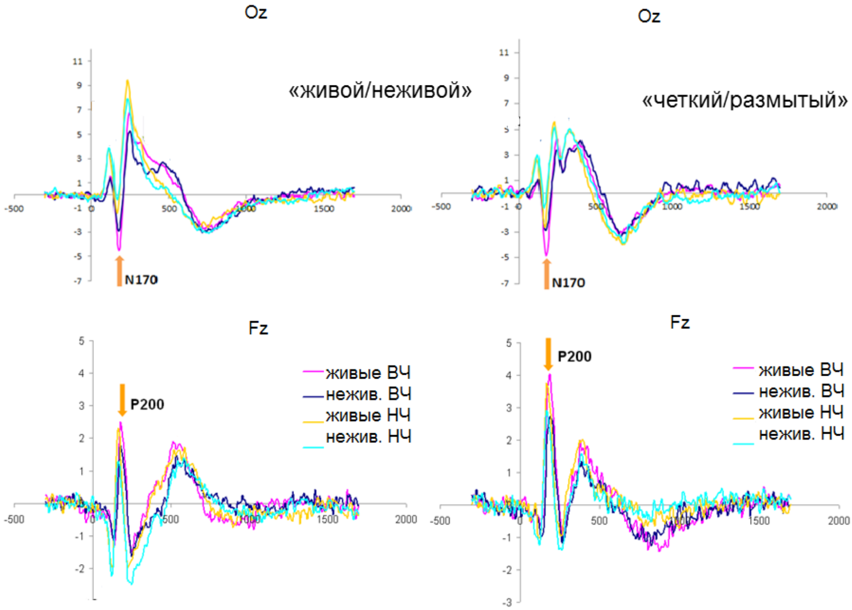


Рис. 3. Примеры вызванных потенциалов при разных инструкциях — слева (живые /неживые) и справа (размытые/четкие). Лобное Fz и затылочное Oz отведение при инструкции наблюдателю определять четкие и размытые, живые и неживые изображения объектов. В лобных отведениях всегда есть отличия только по семантике. Различия по амплитуде вызванных потенциалов только по семантике изображения показаны стрелкой компонент P200. В затылочном отведении Oz есть различия в ответах на физические, и на семантические свойства стимулов. Различия по амплитуде вызванных потенциалов только по семантике изображения показаны стрелкой компонент N170

Давая различную инструкцию наблюдателю, можно выделить и измерить осознанный и неосознанный отклик в мозге. Анализ характера откликов мозга на предъявляемые живые и неживые или на четкие и размытые изображения объектов, позволяет исследователю оценить возможности испытуемого в классификации тестовых изображений. Полученные данные позволяют установить закономерности в решении задач осознаваемой и неосознаваемой классификации наблюдателем при распознавании изображений [Моисеенко и др., 2015, 2016]. Разработанная нами технология пригодна для измере-

ния всех основных параметров механизмов фильтрации и принятия решений, таких как острота зрения, контрастная чувствительность и является новым для офтальмологии, нейроофтальмологии, неврологии, психиатрии и в экспертизе трудоспособности.

* * *

Предложены технологии измерения основных параметров механизмов фильтрации и принятия решений, таких как острота зрения, контрастная чувствительность, которая является новой для офтальмологии, нейроофтальмологии, неврологии, психиатрии и в экспертизе трудоспособности. Новые технологии позволяет обеспечить оценку всех уровней измерения остроты зрения от обнаружения до распознавания и принятия решений. Технология измерений включает предъявление испытуемому визуальных стимулов-оптотипов, различных по пространственно-частотным и семантическим характеристикам, и проведение измерений вызванных потенциалов от затылочных, височных и лобных областей коры. Стимулы отфильтрованы в различных участках видимого диапазона пространственных частот, что позволяет измерять пространственно-частотную контрастную чувствительность в режиме распознавания. Анализ характера вызванных потенциалов на предъявляемые живые/неживые или на четкие/размытые изображения объектов позволяет исследователю оценить возможности испытуемого в классификации тестовых изображений. Полученные результаты предоставляют возможность установить закономерности в решении задач осознаваемой и неосознаваемой классификации наблюдателем при распознавании изображений. По самой высокой пространственной частоте спектра изображения тестового оптотипа, вызывающего отклик мозга, определяют разрешающую способность и, соответственно, остроту зрения испытуемого.

ЛИТЕРАТУРА

Глезер В. Д. Зрение и мышление. Л. Наука, 1993.

Ковальская А. А., Коскин С. А., Бойко Э. В., Шелепин Ю. Е. Объективное исследование остроты зрения методом предпочтительного разглядывания на основе инфракрасной видеоокулографии в целях врачебной экспертизы // Современная оптоме-

трия. 2013. № 2 (62). С. 34–38.

Коскин С.А., Хараузов А.К., Шелепин Ю.Е., Бойко Э.В., Фокин В.А., Труфанов Г.Е., Севостьянов А.В., Пронин С.В., Соболев А.Ф. Современные объективные методы визометрии в целях врачебной экспертизы // Вестник Российской Военно-медицинской академии. 2007. № .4 (20). С. 53–60.

Коскин С.А., Бойко Э.В., Шелепин Ю.Е. Современные методы измерения разрешающей способности зрительной системы. //Оптический журнал. Том. 75, № 1, стр. 22–27, 2008.

Красильников Н.Н. Передача и восприятие изображений, радио и связь. М.1986.

Красильников Н.Н., Шелепин Ю.Е. Функциональная модель зрения. Оптический журнал. 1997; 64 (2): 72–82.

Кропотов Ю.Д. Количественная ЭЭГ, когнитивные вызванные потенциалы и нейротерапия. Донецк, 2010.— 512 с.

Моисеенко Г.А., Шелепин Ю.Е., Хараузов А.К., Пронин С.В., Чихман В.Н., Вахрамеева О.А. Классификация и распознавание изображений живой и неживой природы // Оптический журнал, 82, 10, 2015, стр. 53–64.

Моисеенко Г.А., Вершинина Е.А., Пронин С.В., Чихман В.Н., Михайлова Е.С., Шелепин Ю.Е. Латентные периоды компонентов вызванных потенциалов в задачах классификации изображений, подвергнутых вейвлетной фильтрации //Физиология человека. 2016. Т. 42. № 6. с. 37–48.

Шелепин Ю.Е. Введение в нейроиконику. СПб, 2017.

Шелепин Ю.Е., Колесникова Л.Н., Лекович Ю.И. Визоконтрастометрия. Наука, Л., 1985.

Шелепин Ю.Е., Чихман В.Н., Пронин С.В., Фореман Н. Локальный и глобальный анализ в зрительной системе. В кн.: Современная психофизика. Глава в монографии. М.: Институт психологии РАН; 2009. 542

Авторы

Алексеенко С. В., — д.б.н., ведущий научный сотрудник лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Бондарко В.М., — д.б.н., ведущий научный сотрудник лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Васильев В.Н., — академик РАН, академик РАО, ректор Университета ИТМО, Председатель Совета ректоров вузов Санкт-Петербурга, Вице-президент Российского Союза ректоров.

Васильев П.П., — инженер лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Жукова О.В., — к.псих.н., научный сотрудник лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Кезели А.Р., — д.б.н., профессор, заведующий лабораторией физиологии зрения, Центра экспериментальной биомедицины им. И.С. Бериташвили. Тбилиси, Грузия.

Кирпичников М.П., — академик РАН, зав кафедрой биоинженерии, декан Биологического факультета, проректор, начальник Управления инновационной политики и организации инновационной деятельности МГУ им. М.В. Ломоносова, заведующий лабораторией инженерии

белка Института биоорганический химии им. М.М. Шемякина и Ю.А. Овчинникова РАН.

Ламминпия А.М., — мл. научный сотрудник лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Малахова Е.Ю., — мл. научный сотрудник лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Малашин Р.О., — ст. научный сотрудник, к.т. н., заведующий группой нейронных сетей и искусственного интеллекта Института физиологии им. И.П. Павлова РАН.

Моисеенко Г.А., — мл. научный сотрудник лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Муравьева С.В., — к.м.н., научный сотрудник лаборатории физиологии зрения Института физиологии им. И.П. Павлова РАН.

Островский М.А., — академик РАН, Руководитель секции физиологии Отделения физиологических наук РАН, заведующий кафедрой молекулярной физиологии Биологического факультета МГУ им. М.В. Ломоносова, заведующий лабораторией Института биохимической физики им. Н.М. Эмануэля РАН.

Пономарев С.В., — мл. научный сотрудник сотрудник группы нейронных сетей и искусственного интеллекта Института физиологии им. И.П. Павлова РАН, аспирант Университета ИТМО, Санкт-Петербург.

Пронин С. В. — старший научный сотрудник, лаборатории физиологии зрения, Института физиологии имени И. П.

Соколов А. В., — научный сотрудник отдела радиологии ФГБУ «НМИЦ им. В. А. Алмазова», Санкт-Петербург.

Тибилов А. С., — к. т. н., доцент, Университет ИТМО, Санкт-Петербург.

Труфанов Г. Е., — д. м. н., проф, заведующий кафедрой радиологии ФГБУ «НМИЦ им. В. А. Алмазова», Санкт-Петербург.

Фокин В. А., — д. м. н., проф, заведующий центром радиологии ФГБУ «НМИЦ им. В. А. Алмазова», Санкт-Петербург.

Хараузов А. К., — к. б. н., PhD (Италия), заведующий группой физиологии сенсорных систем приматов Института физиологии им. И. П. Павлова РАН.

Чихман В. Н., — к. т. н., старший научный сотрудник, заведующий лабораторией информационных технологий и математического моделирования Института физиологии им. И. П. Павлова РАН.

Шелепин Е. Ю., — мл. научный сотрудник лаборатории физиологии зрения Института физиологии им. И. П. Павлова РАН.

Шелепин К. Ю., — научный сотрудник лаборатории физиологии речи Института физиологии им. И. П. Павлова РАН, директор ООО «Нейроиконика Ассистив», Санкт-Петербург.

Шелепин Ю. Е., — д. м. н., проф., заведующий лабораторией физиологии зрения Института

физиологии им. И. П. Павлова РАН, профессор Университета ИТМО, профессор СПБГУ.

Якимова Е. Г., — к. б. н., научный сотрудник лаборатории физиологии зрения Института физиологии им. И. П. Павлова РАН.

Научное издание

НЕЙРОТЕХНОЛОГИИ

Под редакцией.
Ю. Е. Шелепина и В. Н. Чихмана

Компьютерная верстка: *Мещерин В.В.*

Подписано в печать 19.11.2018. Формат 60 × 84 $^{1}/_{16}$.
Бумага офсетная. Гарнитура Times. Печать цифровая.
Усл. печ. л. 23,07. Тираж 500 экз. Заказ № 723.

Отпечатано в Издательстве ВВМ .
198095, Санкт-Петербург, ул. Швецова, 41.